Derivatives

Why This Matters

Derivatives measure rate of change. They answer:

> *"How fast is this changing right now?"*

Derivatives are everywhere:

- **Physics**: Velocity, acceleration, forces
- **Economics**: Marginal cost, profit optimization
- **Machine learning**: Gradient descent, backpropagation
- **Engineering**: Control systems, signal processing
- **Data science**: Optimization, curve analysis

Understanding derivatives means understanding **change itself**.

The Big Picture: Instantaneous Rate of Change

Average vs Instantaneous

Average rate of change (slope between two points):

```
         f(b) - f(a)
m_avg = ───────────
           b - a
```

Instantaneous rate of change (slope at one point):

```
          f(x + h) - f(x)
f'(x) = lim ───────────────
        h→0        h
```

This is the derivative.

Visual: Secant to Tangent

```
  Secant line (average):
    y                 •
    |               /
    |             /
    |           /
    |         •
────┴─────────────────▶ x

  Tangent line (instantaneous):
    y
    |         ────
    |       / •
    |     /
────┴─────────────────▶ x
```

As the two points get closer (h → 0), the secant becomes the tangent.

---

1. Definition of the Derivative

The Limit Definition

```
         f(x + h) - f(x)
f'(x) = lim ───────────────
        h→0        h
```

Read as: "f prime of x" (the derivative of f at x)

What it means:

- Change in f divided by change in x
- As the change in x shrinks to zero
- Gives instantaneous rate of change

Alternative Notation

```
f'(x)     (Lagrange notation)
df/dx     (Leibniz notation)
dy/dx     (if y = f(x))
Df(x)     (Operator notation)
```

All mean the same thing.

---

2. Computing Derivatives from the Definition

Example 1: f(x) = x²

```
         (x+h)² - x²
f'(x) = lim ──────────
        h→0      h

        x² + 2xh + h² - x²
      = lim ─────────────────
        h→0         h

        2xh + h²
      = lim ────────
        h→0     h

      = lim (2x + h)
        h→0

      = 2x
```

Result: If $f(x) = x^2$, then $f'(x) = 2x$

Example 2: f(x) = 3x + 1

```
         (3(x+h) + 1) - (3x + 1)
f'(x) = lim ────────────────────
        h→0          h

        3x + 3h + 1 - 3x - 1
      = lim ───────────────
        h→0          h

        3h
      = lim ──
        h→0  h

      = 3
```

Result: If f(x) = 3x + 1, then f'(x) = 3

General: Linear functions have constant derivatives (their slopes).

---

3. Common Derivative Formulas

Instead of using the limit definition every time, we have formulas:

Power Rule

```
d/dx(xⁿ) = nxⁿ⁻¹
```

Examples:

```
d/dx(x³) = 3x²
d/dx(x⁵) = 5x⁴
d/dx(x) = 1x⁰ = 1
d/dx(1) = 0  (constant)
```

Constant Rule

```
d/dx(c) = 0

The derivative of a constant is zero (no change)
```

Constant Multiple Rule

```
d/dx(cf(x)) = c·f'(x)

Constants pull out
```

Example:

```
d/dx(5x³) = 5·3x² = 15x²
```

Sum/Difference Rule

```
d/dx(f(x) + g(x)) = f'(x) + g'(x)

Derivatives distribute over addition
```

Example:

```
d/dx(x³ + 2x² - 5x + 7)
= 3x² + 4x - 5 + 0
= 3x² + 4x - 5
```

---

4. Product and Quotient Rules

Product Rule

```
d/dx(f·g) = f'·g + f·g'
```

NOT f'·g'!

Example:

```
d/dx(x²·sin(x))

f = x²,  f' = 2x
g = sin(x), g' = cos(x)

= 2x·sin(x) + x²·cos(x)
```

Quotient Rule

```
       f        f'·g - f·g'
d/dx( ─ ) = ─────────────
       g           g²
```

"Low d-high minus high d-low, all over low squared"

Example:

```
       x²
d/dx( ──── )
       x+1

f = x², f' = 2x
g = x+1, g' = 1

  2x(x+1) - x²(1)     2x² + 2x - x²     x² + 2x
```

```
= ─────────────── = ─────────────── = ───────────
      (x+1)²             (x+1)²           (x+1)²
```

---

5. Chain Rule (The Most Important)

The Rule

For composite functions:

```
d/dx(f(g(x))) = f'(g(x))·g'(x)
```

In words: "Derivative of outer × derivative of inner"

Example 1: $(x^2 + 1)^3$

```
Outer function: u³
Inner function: u = x² + 1

d/dx((x²+1)³) = 3(x²+1)²·(2x)
              = 6x(x²+1)²
```

Example 2: $\sin(x^2)$

```
Outer: sin(u)
Inner: u = x²

d/dx(sin(x²)) = cos(x²)·(2x)
              = 2x·cos(x²)
```

Example 3: e^{x^2}

```
d/dx(eˣ²) = eˣ²·(2x)
          = 2x·eˣ²
```

Why It Matters

Most complex derivatives need the chain rule.

```
// In neural networks, backpropagation is repeated chain rule
function backprop(layers) {
  let gradient = 1;
  for (let i = layers.length - 1; i >= 0; i--) {
    gradient *= layers[i].derivative();  // Chain rule!
  }
  return gradient;
}
```

---

6. Derivatives of Standard Functions

Polynomials

$$\frac{d}{dx}(x^n) = nx^{n-1}$$

Exponential

$$\frac{d}{dx}(e^x) = e^x \quad \text{(special property!)}$$

$$\frac{d}{dx}(a^x) = a^x \cdot \ln(a)$$

Logarithmic

$$\frac{d}{dx}(\ln(x)) = 1/x$$

$$\frac{d}{dx}(\log_a(x)) = 1/(x \cdot \ln(a))$$

Trigonometric

$$\frac{d}{dx}(\sin(x)) = \cos(x)$$
$$\frac{d}{dx}(\cos(x)) = -\sin(x)$$
$$\frac{d}{dx}(\tan(x)) = \sec^2(x) = 1/\cos^2(x)$$

Inverse Trig

$$\frac{d}{dx}(\sin^{-1}(x)) = 1/\sqrt{1-x^2}$$
$$\frac{d}{dx}(\cos^{-1}(x)) = -1/\sqrt{1-x^2}$$
$$\frac{d}{dx}(\tan^{-1}(x)) = 1/(1+x^2)$$

---

7. What Derivatives Tell Us

Slope of the Tangent Line

At any point, f'(x) is the slope of the tangent line.

```
     y
     │      ⁄
     │    ⁄ tangent (slope = f'(a))
     │  ⁄
     │●
─────┴──────────▶ x
     a
```

Tangent line equation at x = a:

$$y - f(a) = f'(a)(x - a)$$

$$\text{or: } y = f'(a)(x - a) + f(a)$$

Increasing/Decreasing

```
f'(x) > 0  →  f is increasing
f'(x) < 0  →  f is decreasing
f'(x) = 0  →  f has a horizontal tangent (critical point)
```

```
    |    /\    f' > 0: rising
    |   /  \   f' = 0: peak
    |  /    \  f' < 0: falling
____|____________
```

Critical Points

Where f'(x) = 0 or f'(x) doesn't exist.

These are **candidates** for max/min values.

Example: $f(x) = x^2 - 4x + 3$

```
f'(x) = 2x - 4 = 0
x = 2  (critical point)

f(2) = 4 - 8 + 3 = -1  (minimum)
```

Concavity (Second Derivative)

Second derivative f''(x) = derivative of f'(x)

```
f''(x) > 0  →  concave up (∪)
f''(x) < 0  →  concave down (∩)
f''(x) = 0  →  possible inflection point
```

---

8. Applications

Velocity and Acceleration

Position function: s(t) **Velocity**: v(t) = s'(t) **Acceleration**: a(t) = v'(t) = s''(t)

Example: $s(t) = -16t^2 + 64t + 5$

```
Velocity: v(t) = -32t + 64
Acceleration: a(t) = -32 ft/s²  (gravity)

At t = 2:
v(2) = -32(2) + 64 = 0  (peak height)
```

Optimization

Find maximum or minimum values.

Method:

1. Find f'(x)
2. Solve f'(x) = 0 for critical points
3. Test which is max/min (using second derivative or endpoints)

Example: Maximize area of rectangle with perimeter 100

```
Let width = x, then height = (100-2x)/2 = 50-x

Area: A(x) = x(50-x) = 50x - x²

A'(x) = 50 - 2x = 0
x = 25

Max area = 25(25) = 625 sq units (square shape)
```

Marginal Analysis (Economics)

Marginal cost = derivative of cost function **Marginal revenue** = derivative of revenue function **Marginal profit** = derivative of profit function

```
C(x) = 1000 + 5x + 0.01x²
C'(x) = 5 + 0.02x  (marginal cost)

At x = 100: C'(100) = 5 + 2 = $7 per unit
```

Machine Learning: Gradient Descent

Update rule:

```
θ_new = θ_old - α·∇J(θ)
        ↑       ↑    ↑
      param  learn gradient (derivative!)
             rate
```

In code:

```
function gradientDescent(f, df, x0, learningRate, iterations) {
  let x = x0;
  for (let i = 0; i < iterations; i++) {
    x = x - learningRate * df(x);  // Move opposite to gradient
  }
  return x;
}

// Minimize f(x) = x²
const f = x => x**2;
const df = x => 2*x;
gradientDescent(f, df, 10, 0.1, 100);  // Converges to 0
```

Related Rates

When two quantities change over time, relate their derivatives.

Example: Balloon radius increasing at 2 cm/s. How fast is volume increasing?

```
V = (4/3)πr³

dV/dt = 4πr²·dr/dt  (chain rule)

If dr/dt = 2 cm/s and r = 5 cm:
dV/dt = 4π(25)(2) = 200π cm³/s
```

---

9. Programming Derivatives

Numerical Approximation

```
function derivative(f, x, h = 1e-5) {
  return (f(x + h) - f(x - h)) / (2*h);
}

// Example: f(x) = x³
const f = x => x**3;
derivative(f, 2);  // ≈ 12 (exact: 3(2²) = 12)
```

Automatic Differentiation (Modern ML)

```
// Simplified dual number (stores value and derivative)
class Dual {
  constructor(value, derivative = 0) {
    this.value = value;
    this.derivative = derivative;
  }

  add(other) {
    return new Dual(
      this.value + other.value,
      this.derivative + other.derivative
    );
  }

  multiply(other) {
    return new Dual(
      this.value * other.value,
      this.derivative * other.value + this.value * other.derivative
    );
  }
}

// Compute f(x) = x² at x=3
const x = new Dual(3, 1);  // x=3, dx/dx=1
const result = x.multiply(x);
```

```
console.log(result.value);      // 9
console.log(result.derivative); // 6 (exact: 2x = 6)
```

---

10. Higher-Order Derivatives

Notation

```
f'(x)   = first derivative
f''(x)  = second derivative (derivative of f')
f'''(x) = third derivative
fⁿ⁾(x) = nth derivative
```

Leibniz notation:

```
dy/dx, d²y/dx², d³y/dx³, ...
```

Example

```
f(x) = x⁴

f'(x) = 4x³
f''(x) = 12x²
f'''(x) = 24x
f⁽⁴⁾(x) = 24
f⁽⁵⁾(x) = 0  (all higher derivatives are zero)
```

Physical Meaning

```
Position:     s(t)
Velocity:     v(t) = s'(t)
Acceleration: a(t) = s''(t)
Jerk:         j(t) = s'''(t)  (rate of change of acceleration)
```

Common Mistakes & Misconceptions

❌ "Derivative of f·g is f'·g'"

No! Use product rule: f'·g + f·g'

❌ "d/dx(f/g) = f'/g'"

No! Use quotient rule: (f'·g - f·g')/g^2

❌ "Forgetting the chain rule"

```
d/dx(sin(x²)) ≠ cos(x²)

Correct: cos(x²)·2x
```

❌ "f'(x) = 0 means maximum"

Could be minimum, or neither (inflection point). Must test.

❌ "Derivative doesn't exist = function doesn't exist"

Function can exist but not be differentiable (sharp corner, vertical tangent).

Tiny Practice

Find derivatives:

1. $f(x) = x^3 - 2x + 5$
2. $f(x) = 3x^4 + 2x^2 - 7$
3. $f(x) = (x^2 + 1)(x - 2)$
4. $f(x) = x^2/x+1$
5. $f(x) = (x^2 + 1)^3$

Applications: 6. If $s(t) = -16t^2 + 32t$, find velocity at t = 1 7. Find critical points of $f(x) = x^3 - 3x$ 8. At what x does $f(x) = x^2 - 4x + 3$ have minimum?

▶ Answers

Summary Cheat Sheet

Definition

```
        f(x+h) - f(x)
f'(x) = lim ─────────────
       h→0       h
```

Key Rules

Function	Derivative
x^n	nx^{n-1}
e^x	e^x
ln(x)	1/x
sin(x)	cos(x)
cos(x)	-sin(x)
cf(x)	cf'(x)
f + g	f' + g'
fg	f'g + fg'
f/g	$(f'g - fg')/g^2$
f(g(x))	f'(g(x))·g'(x)

Interpretation

```
f'(x) > 0  →  increasing
f'(x) < 0  →  decreasing
f'(x) = 0  →  critical point

f''(x) > 0 →  concave up
f''(x) < 0 →  concave down
```

Applications

```
Velocity: v = ds/dt
Acceleration: a = dv/dt
Optimization: Set f'(x) = 0, solve
Tangent line: y = f'(a)(x-a) + f(a)
```

Programming

```
// Numerical
const df = (f, x, h=1e-5) => (f(x+h) - f(x-h))/(2*h);

// Chain rule example
const d_sin_x2 = x => Math.cos(x*x) * 2*x;
```

Next Steps

Derivatives measure instantaneous rate of change. You now understand:

- The limit definition
- Common derivative rules
- Applications to optimization and motion

Next and finally, we'll explore **Integrals**—the reverse of derivatives, measuring accumulation.

Continue to: 12-integrals.md