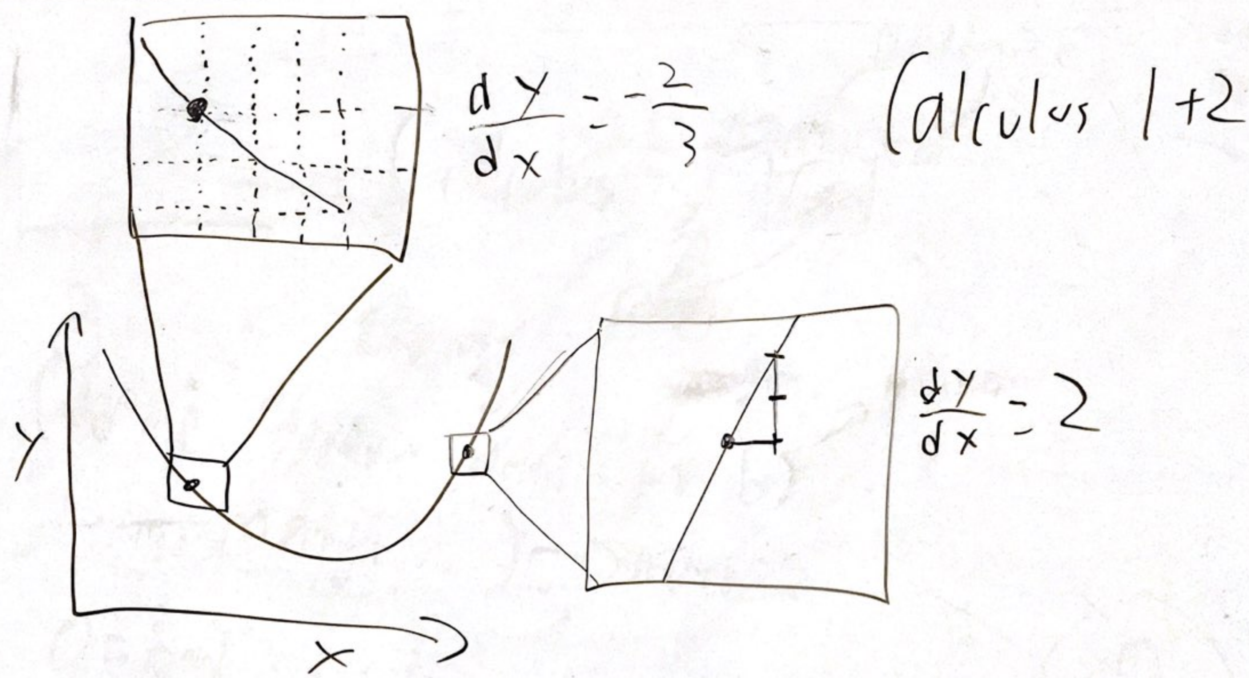
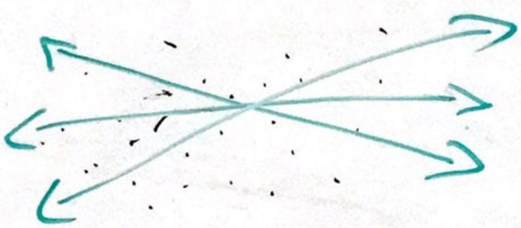
$$z = x \cdot y \rightarrow \frac{\partial w}{\partial z} = \frac{\partial w}{\partial x} \cdot y + \frac{\partial w}{\partial y} \cdot x$$

$$z = x + y \rightarrow \frac{\partial w}{\partial z} = \frac{\partial w}{\partial x} + \frac{\partial w}{\partial y}$$

chain rule: $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x}$



$\frac{dy}{dx} = -\frac{2}{3}$
Calculus 1+2
$\frac{dy}{dx} = 2$
y
x



Optimize

$Y = Mx + b$

const

List posn

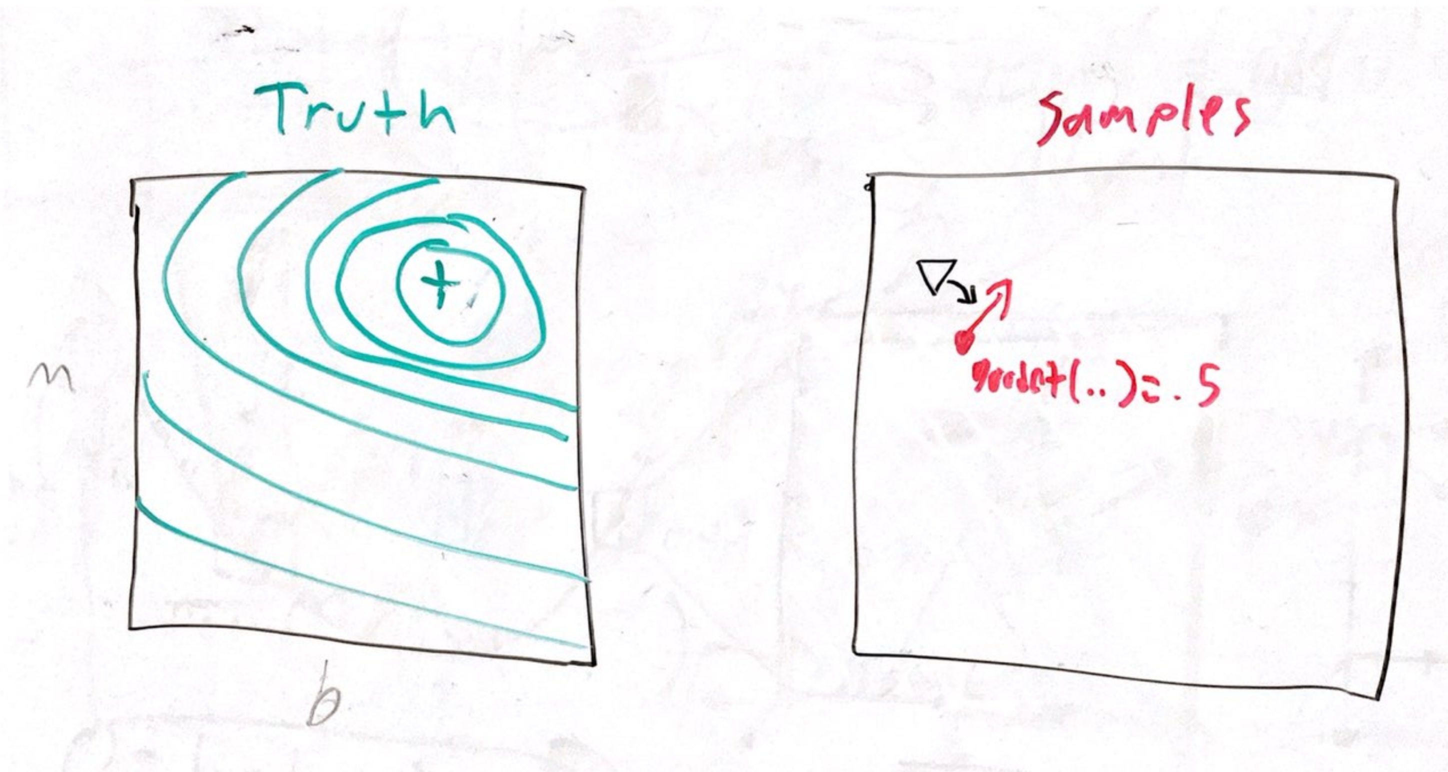
number

number

(define (goodfit data m b)

... math ...) → number

0 = bad ... 1 = perfect


Truth
Samples
m
b
∇


Truth

Calculate Gradient

1) $x^2 \longrightarrow 2x$ — Exact answer
- control flow?
- functions?

2) $\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$ — Black Box function
- ERR: DIV BY 0
- Values of ϵ
- Many input vars?

3) Wengert 1964

dual numbers: $\overset{\text{vec}(2)}{\left[x, \frac{\partial x}{\partial \boxed{?}}\right]}$

```
def multiply(M, V):
    [M·V, M(1)·V(2) + M(2)·V(1)]
```

$f(x) = x^2 + 5$

```
def f(x: vec[2]) -> vec[2]:
    let x_sqr = x^2
    return x_sqr + 5
```

At $x=3$

$x=3$ $\frac{\partial x}{\partial x} = 1$

$x=9$ $\frac{\partial x\_sqr}{\partial x} = 6$

$x=14$ $\frac{\partial f(x)}{\partial x} = 6$

```
def foo(x: dual, y: dual):
    return x·y
```

$x = 3,\ y = 4$

■ = Runtime

Wengert

①

initial values $x = [x_{val}, \frac{\partial x}{\partial x}]$ $y = [y_{val}, \frac{\partial y}{\partial x}]$

$= [3, 1]$ $= [4, 0]$

returned value $x \cdot y = [xy_{val}, \frac{\partial xy}{\partial x}]$

$= [12, 4]$

②

$x = [x_{val}, \frac{\partial x}{\partial y}]$ $y = [y_{val}, \frac{\partial y}{\partial y}]$

$= [3, 0]$ $= [4, 1]$

$xy = [xy_{val}, \frac{\partial xy}{\partial y}]$

$= [12, 3]$

$\nabla foo(3,4) = (4,3)$

Rall

$x = \begin{Bmatrix} f: x_{val} \\ df: [\frac{\partial x}{\partial x}, \frac{\partial x}{\partial y}] \end{Bmatrix}$ $y = \begin{Bmatrix} f: x_{val} \\ df: [\frac{\partial y}{\partial x}, \frac{\partial y}{\partial y}] \end{Bmatrix}$

$= \begin{Bmatrix} f: 3 \\ df: [1,0] \end{Bmatrix}$ $= \begin{Bmatrix} f: 4 \\ df: [0,1] \end{Bmatrix}$

$xy = \begin{Bmatrix} f: xy_{val} \\ df: [\frac{\partial xy}{\partial x}, \frac{\partial xy}{\partial y}] \end{Bmatrix}$

$= \begin{Bmatrix} f: 12 \\ df: [4,3] \end{Bmatrix}$

JAKE

```
def foo(x, y)
  CONSTRUCT D(b)/D(x) ->
  let a = x+5
  let b = a·y
```

```
def foo(x, y)
  let dxdx = 1
  let dydx = 1
  let a = x+5
  let dadx = dxdx
  let b = a·y
  let dbdx = dadx·y + a·dydx
```

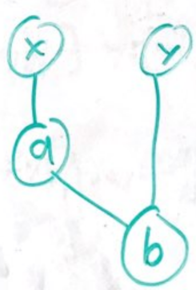
Reverse Mode

def foo (x, y)

CONSTRUCT D(b)/D(x and y) →

let a = x + 5

let b = a · y



def foo(x, y) $x = 2,\ y = 3$

let a = x + 5 $a = 7$

let b = a · y $b = 21$

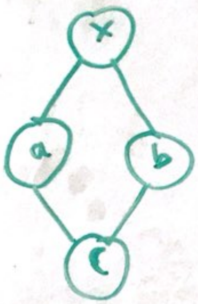
let dbdb = 1 $\frac{\partial b}{\partial b} = 1$

let dbda = y · dbdb $\frac{\partial b}{\partial a} = 3$

let dbdy = a · dbdb $\frac{\partial b}{\partial y} = 7$

let dbdx = dbda $\frac{\partial b}{\partial x} = 3$

$\nabla foo(x, y) = [3, 7]$



def $f(x)$

$a = 2x$

$b = x^2$

$c = a + b$

dcdb $= 1$

dcda $= 1$

dcdx $= \frac{dc}{db} \cdot \frac{db}{dx} =$ dcdb $\cdot 2x$ $\rightarrow$ dcdx$_1$ = dcdb $\cdot 2x$

dcdx$_2$ = dcda $\cdot 2$

dcdx = dcdx$_1$ + dcdx$_2$

dcdx $\mathrel{+}= \frac{dc}{da} \cdot \frac{da}{dx} =$ dcda $\cdot 2$

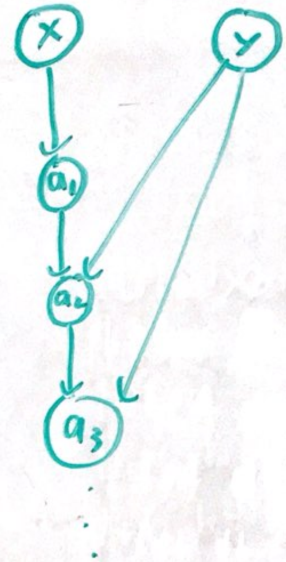
$\rightarrow$ $x_1 = x_2 = x$

$a = 2x_1$

$b = 2x_2^2$

$$\frac{dc}{dx} = \frac{dc}{dx_1} + \frac{dc}{dx_2}$$

```
def foo(x,y):
    accum = x
    for i in 1..10:
        accum = accum * y
```



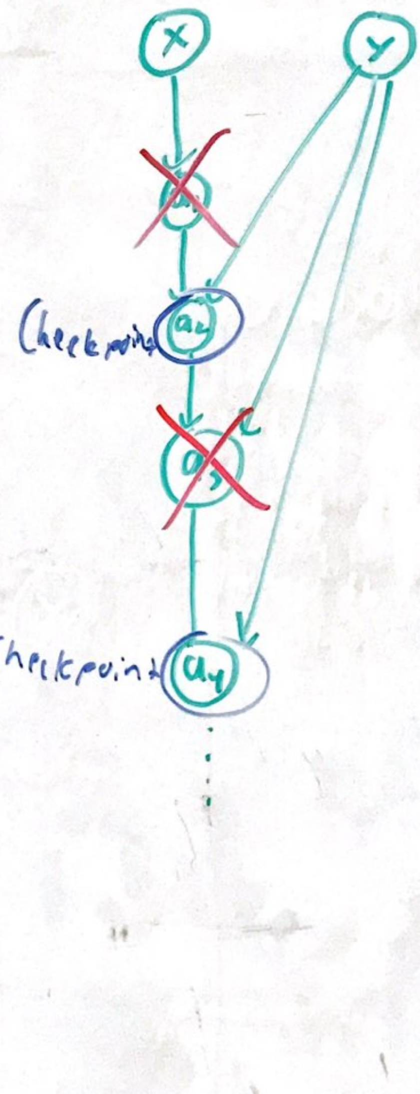
Forward Mode: $O(1)$ Mem

Reverse Mode: $O(c)$ Mem

↑

Computational complexity

```
def foo(x, y)
    a1 = x
    a2 = a1 · y
    free a1
    a3 = a2 · y
    a4 = a3 · y
    free a3
    ⋮
    da4da10 = ....
    a3* = a2 · y
    da3da10 = da4da10 · y
    dyda10 += da4da10 · a3*
    free a3*
    da2da10 = da3da10 · y
    dyda10 += da3da10 · a2
    a1* = x
    ....
```



- Griewank 1996: ADOL-C
  Reverse mode everywhere

- Recently: GPGPU

- Torchscript: Python -> AD-focused IR

Extra time: 2nd derivative

```
def f(x)
    a = x^3
    b = 5a
    ret b
```

Symbolic:

$b = 5x^3$

$\frac{\partial b}{\partial x} = 15x^2$

$\frac{\partial^2 b}{\partial x^2} = 30x$

$\rightarrow$

$a = x^3$

$b = 5a$

dbdb = 1

dbda = 5

dbdx = $\frac{db}{da} \cdot \frac{da}{dx}$ = dbda $\cdot 3x^2$

$\rightarrow$

$a = x^3$

$b = 5a$

dbdb = 1

dbda = 5

dbdx = dbda $\cdot 3x^2$

∂dbdx∂dbdx = 1

∂dbdx∂dbda = $3x^2$

∂dbdx∂x = $6 \cdot x \cdot$ dbda

$x = 2$

$a = 8$

$b = 40 = 5x^3$

dbdb = 1

dbda = 5

dbdx = $60 = 15x^2$

$\frac{\partial \frac{db}{dx}}{\partial \frac{db}{dx}} = 1$

$\frac{\partial \frac{db}{dx}}{\partial \frac{db}{da}} = 12$

$\frac{\partial \frac{db}{dx}}{\partial x} = 60 = 30x$