Automatic Differentiation

CSC412/2506
Winter 2018

Slides based on the excellent review:
Baydin, A. G., Pearlmutter, B. A., Radul, A. A., & Siskind, J. M. (2015). *Automatic differentiation in machine learning: a survey.*
http://arxiv.org/abs/1502.05767

What is AD?

“A family of techniques similar to but more general than back propagation for efficiently and accurately evaluating derivatives of numeric functions expressed as computer programs.”

All numerical computations are composed of a finite set of elementary operations.
These elementary operations have known derivatives.
Systematically apply the chain rule of differential calculus.

4 Categories of Derivatives for Computer Programs

1. Manual Differentiation

   (computing by hand and coding the result)

2. Numerical Differentiation

   (e.g. finite differences approx.)

3. Symbolic Differentiation

   (Mathematica, Maple…)

4. Automatic Differentiation

   (subject of this tutorial)

Why do we need AD?

Manual Differentiation is time consuming and error prone.

Numerical Differentiation scales poorly and highly susceptible to roundoff/truncation errors.

Symbolic Differentiation ‘swells’ quickly as derivative expressions become very complex.

Also, both Manual and Symbolic require closed-form mathematical expression.

$$l_1 = x$$
$$l_{n+1} = 4l_n(1 - l_n)$$

$$f(x) = l_4 = 64x(1-x)(1-2x)^2(1-8x+8x^2)^2$$

Manual Differentiation

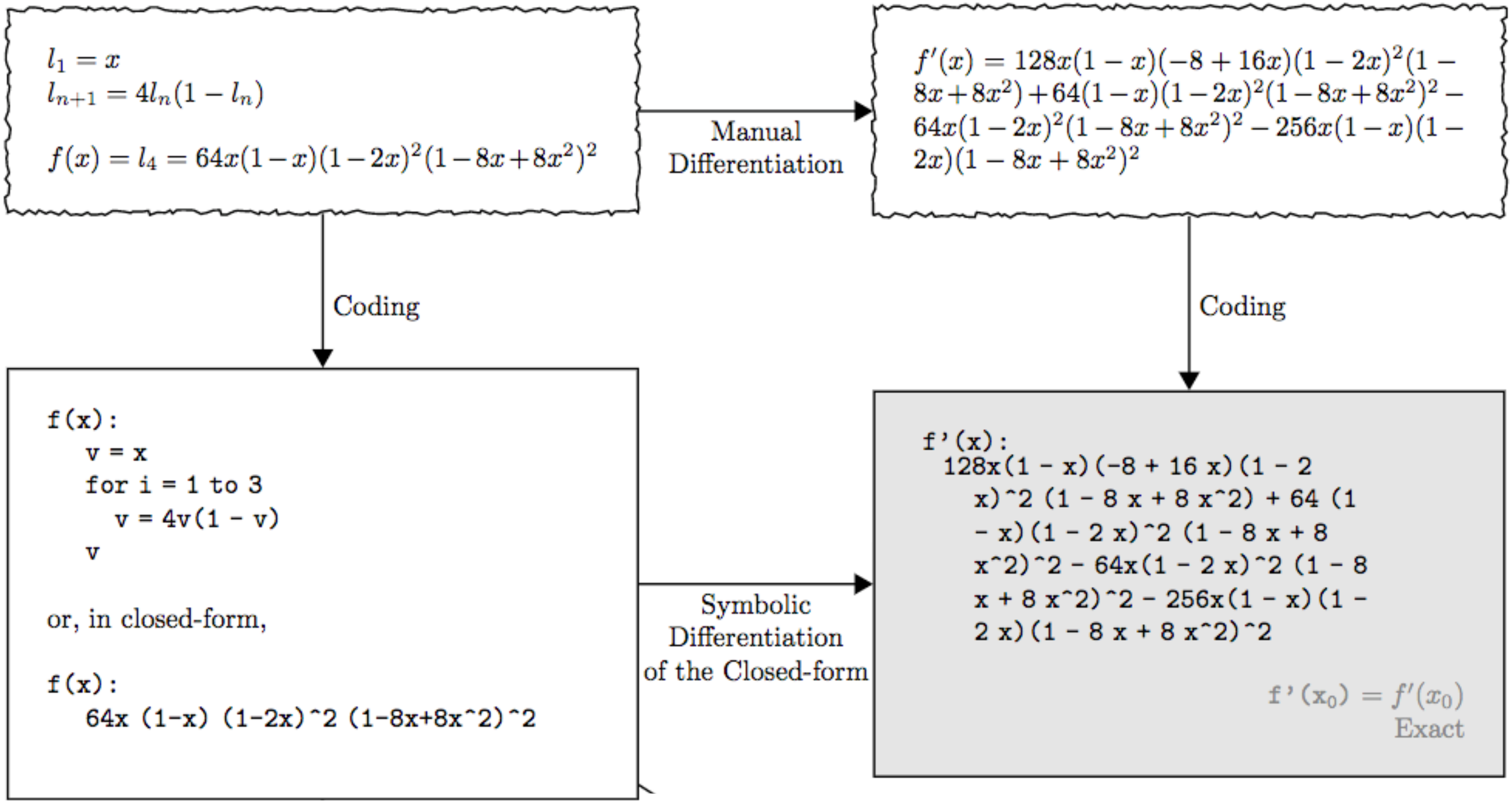
$$f'(x) = 128x(1-x)(-8+16x)(1-2x)^2(1-8x+8x^2)+64(1-x)(1-2x)^2(1-8x+8x^2)^2-64x(1-2x)^2(1-8x+8x^2)^2-256x(1-x)(1-2x)(1-8x+8x^2)^2$$

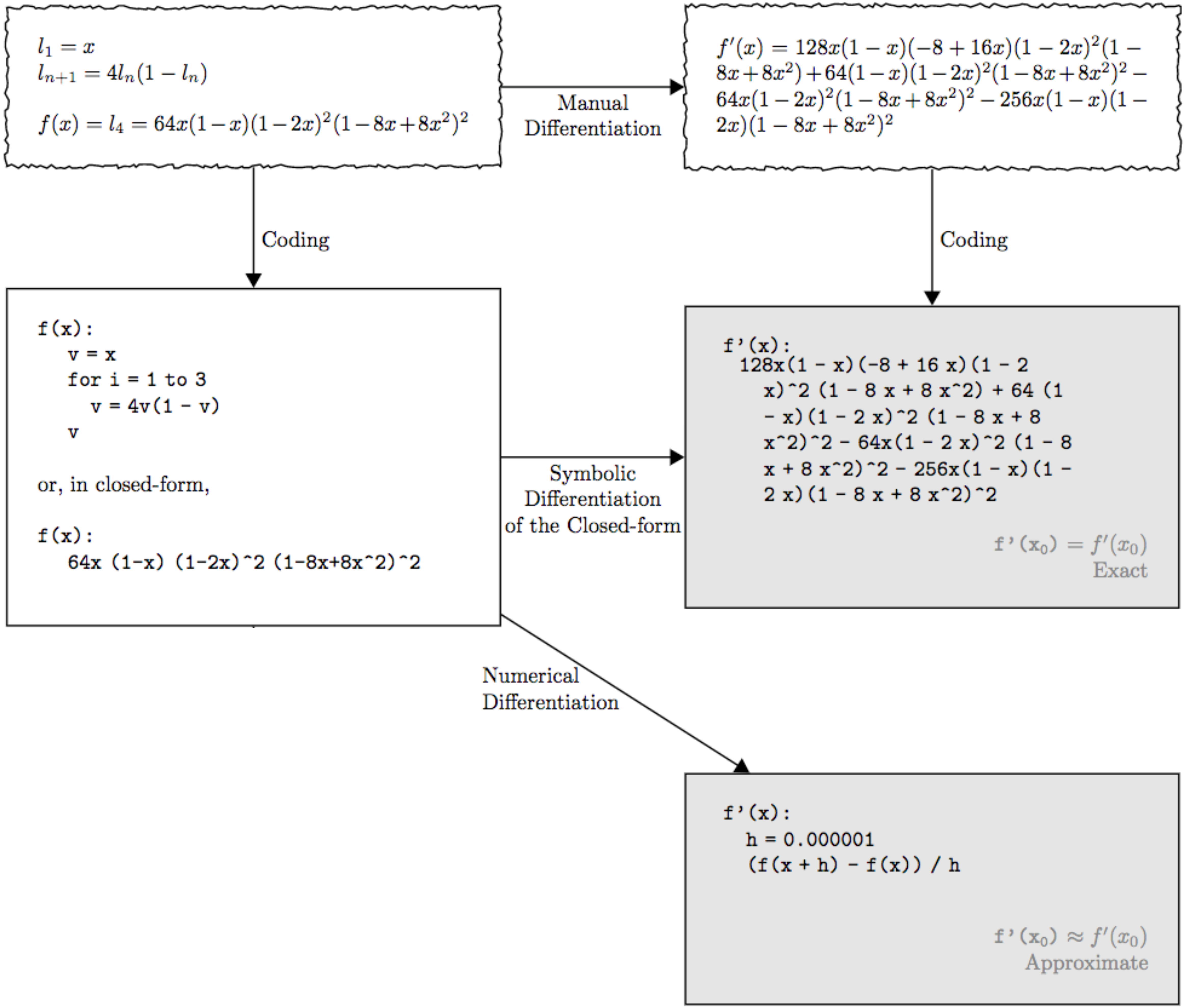
Coding

```
f(x):
  v = x
  for i = 1 to 3
    v = 4v(1 - v)
  v
```

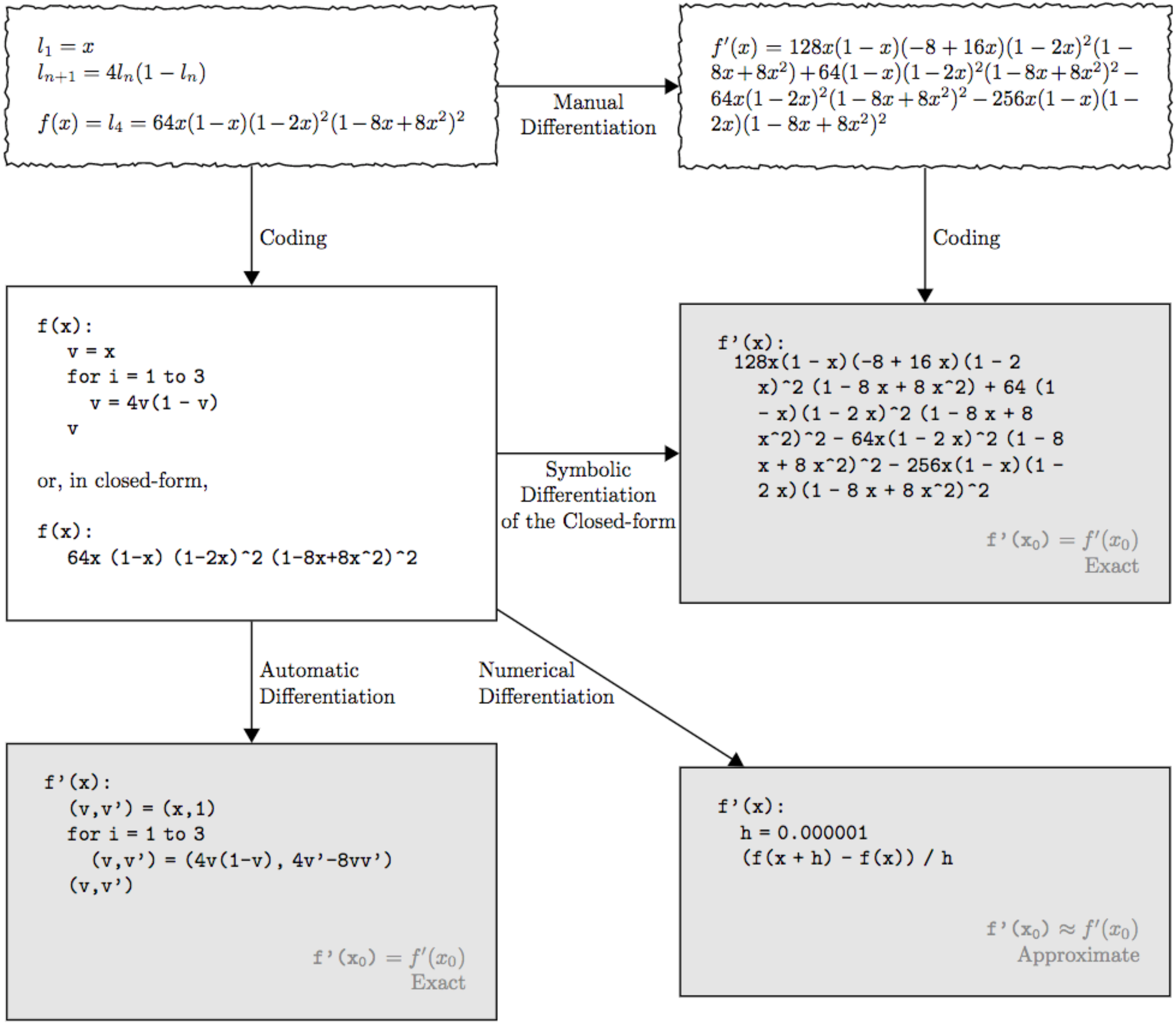
or, in closed-form,

```
f(x):
  64x (1-x) (1-2x)^2 (1-8x+8x^2)^2
```


$l_1 = x$
$l_{n+1} = 4l_n(1 - l_n)$
$f(x) = l_4 = 64x(1-x)(1-2x)^2(1-8x+8x^2)^2$
Manual Differentiation
$f'(x) = 128x(1-x)(-8+16x)(1-2x)^2(1-8x+8x^2) + 64(1-x)(1-2x)^2(1-8x+8x^2)^2 - 64x(1-2x)^2(1-8x+8x^2)^2 - 256x(1-x)(1-2x)(1-8x+8x^2)^2$
Coding
Coding
f(x):
v = x
for i = 1 to 3
v = 4v(1 - v)
v
or, in closed-form,
f(x):
64x (1-x) (1-2x)^2 (1-8x+8x^2)^2
Symbolic Differentiation of the Closed-form
f'(x):
128x(1 - x)(-8 + 16 x)(1 - 2 x)^2 (1 - 8 x + 8 x^2) + 64 (1 - x)(1 - 2 x)^2 (1 - 8 x + 8 x^2)^2 - 64x(1 - 2 x)^2 (1 - 8 x + 8 x^2)^2 - 256x(1 - x)(1 - 2 x)(1 - 8 x + 8 x^2)^2
f'(x_0) = $f'(x_0)$
Exact



$l_1 = x$
$l_{n+1} = 4l_n(1 - l_n)$
$f(x) = l_4 = 64x(1-x)(1-2x)^2(1-8x+8x^2)^2$
Manual Differentiation
$f'(x) = 128x(1-x)(-8+16x)(1-2x)^2(1-8x+8x^2) + 64(1-x)(1-2x)^2(1-8x+8x^2)^2 - 64x(1-2x)^2(1-8x+8x^2)^2 - 256x(1-x)(1-2x)(1-8x+8x^2)^2$
Coding
Coding
f(x):
v = x
for i = 1 to 3
v = 4v(1 - v)
v
or, in closed-form,
f(x):
64x (1-x) (1-2x)^2 (1-8x+8x^2)^2
Symbolic Differentiation of the Closed-form
f'(x):
128x(1 - x)(-8 + 16 x)(1 - 2 x)^2 (1 - 8 x + 8 x^2) + 64 (1 - x)(1 - 2 x)^2 (1 - 8 x + 8 x^2)^2 - 64x(1 - 2 x)^2 (1 - 8 x + 8 x^2)^2 - 256x(1 - x)(1 - 2 x)(1 - 8 x + 8 x^2)^2
f'(x$_0$) = $f'(x_0)$
Exact
Numerical Differentiation
f'(x):
h = 0.000001
(f(x + h) - f(x)) / h
f'(x$_0$) $\approx f'(x_0)$
Approximate



$l_1 = x$
$l_{n+1} = 4l_n(1 - l_n)$
$f(x) = l_4 = 64x(1-x)(1-2x)^2(1-8x+8x^2)^2$
Manual Differentiation
$f'(x) = 128x(1-x)(-8+16x)(1-2x)^2(1-8x+8x^2)+64(1-x)(1-2x)^2(1-8x+8x^2)^2-64x(1-2x)^2(1-8x+8x^2)^2-256x(1-x)(1-2x)(1-8x+8x^2)^2$
Coding
Coding
f(x):
v = x
for i = 1 to 3
v = 4v(1 - v)
v
or, in closed-form,
f(x):
64x (1-x) (1-2x)^2 (1-8x+8x^2)^2
Symbolic Differentiation of the Closed-form
f'(x):
128x(1 - x)(-8 + 16 x)(1 - 2 x)^2 (1 - 8 x + 8 x^2) + 64 (1 - x)(1 - 2 x)^2 (1 - 8 x + 8 x^2)^2 - 64x(1 - 2 x)^2 (1 - 8 x + 8 x^2)^2 - 256x(1 - x)(1 - 2 x)(1 - 8 x + 8 x^2)^2
f'(x$_0$) = $f'(x_0)$
Exact
Automatic Differentiation
Numerical Differentiation
f'(x):
(v,v') = (x,1)
for i = 1 to 3
(v,v') = (4v(1-v), 4v'-8vv')
(v,v')
f'(x$_0$) = $f'(x_0)$
Exact
f'(x):
h = 0.000001
(f(x + h) - f(x)) / h
f'(x$_0$) ≈ $f'(x_0)$
Approximate

What is Automatic Differentiation?

2 Modes of AD

$$y = f(g(h(x))) = f(g(h(w_0))) = f(g(w_1)) = f(w_2) = w_3$$

$$\frac{dy}{dx} = \boxed{\frac{dy}{dw_2}\boxed{\frac{dw_2}{dw_1}\boxed{\frac{dw_1}{dx}}}}$$

Forward Accumulation Mode: chain rule inside to outside

$$\boxed{dw_1/dx} \longrightarrow \boxed{dw_2/dx} \longrightarrow \boxed{dy/dx}$$

2 Modes of AD

$$y = f(g(h(x))) = f(g(h(w_0))) = f(g(w_1)) = f(w_2) = w_3$$

$$\frac{dy}{dx} = \boxed{\boxed{\boxed{\frac{dy}{dw_2}}\frac{dw_2}{dw_1}}\frac{dw_1}{dx}}$$

Reverse Accumulation Mode: chain rule outside to inside

$$\boxed{dy/dw_2} \longrightarrow \boxed{dy/dw_1} \longrightarrow \boxed{dy/dx}$$

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace	Forward Tangent (Derivative) Trace
↓	↓

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace		Forward Tangent (Derivative) Trace	
$v_{-1} = x_1$	$= 2$	$\dot{v}_{-1} = \dot{x}_1$	$= 1$
$v_0 = x_2$	$= 5$	$\dot{v}_0 = \dot{x}_2$	$= 0$

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace			Forward Tangent (Derivative) Trace		
v_{-1}	$= x_1$	$= 2$	$\dot{v}_{-1}$	$= \dot{x}_1$	$= 1$
v_0	$= x_2$	$= 5$	$\dot{v}_0$	$= \dot{x}_2$	$= 0$
v_1	$= \ln v_{-1}$	$= \ln 2$	$\dot{v}_1$	$= \dot{v}_{-1}/v_{-1}$	$= 1/2$

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$

Forward Tangent (Derivative) Trace		
$\dot{v}_{-1}$	$= \dot{x}_1$	$= 1$
$\dot{v}_0$	$= \dot{x}_2$	$= 0$
$\dot{v}_1$	$= \dot{v}_{-1}/v_{-1}$	$= 1/2$
$\dot{v}_2?$		

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace			Forward Tangent (Derivative) Trace		
v_{-1}	$= x_1$	$= 2$	$\dot{v}_{-1}$	$= \dot{x}_1$	$= 1$
v_0	$= x_2$	$= 5$	$\dot{v}_0$	$= \dot{x}_2$	$= 0$
v_1	$= \ln v_{-1}$	$= \ln 2$	$\dot{v}_1$	$= \dot{v}_{-1}/v_{-1}$	$= 1/2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$	$\dot{v}_2$	$= \dot{v}_{-1} \times v_0 + \dot{v}_0 \times v_{-1}$	$= 1 \times 5 + 0 \times 2$

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$

Forward Tangent (Derivative) Trace		
$\dot{v}_{-1}$	$= \dot{x}_1$	$= 1$
$\dot{v}_0$	$= \dot{x}_2$	$= 0$
$\dot{v}_1$	$= \dot{v}_{-1}/v_{-1}$	$= 1/2$
$\dot{v}_2$	$= \dot{v}_{-1} \times v_0 + \dot{v}_0 \times v_{-1}$	$= 1 \times 5 + 0 \times 2$
$\dot{v}_3?$		

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$

Forward Tangent (Derivative) Trace		
$\dot{v}_{-1}$	$= \dot{x}_1$	$= 1$
$\dot{v}_0$	$= \dot{x}_2$	$= 0$
$\dot{v}_1$	$= \dot{v}_{-1}/v_{-1}$	$= 1/2$
$\dot{v}_2$	$= \dot{v}_{-1} \times v_0 + \dot{v}_0 \times v_{-1}$	$= 1 \times 5 + 0 \times 2$
$\dot{v}_3$	$= \dot{v}_0 \times \cos v_0$	$= 0 \times \cos 5$

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace

v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$

Forward Tangent (Derivative) Trace

$\dot{v}_{-1}$	$= \dot{x}_1$	$= 1$
$\dot{v}_0$	$= \dot{x}_2$	$= 0$
$\dot{v}_1$	$= \dot{v}_{-1}/v_{-1}$	$= 1/2$
$\dot{v}_2$	$= \dot{v}_{-1} \times v_0 + \dot{v}_0 \times v_{-1}$	$= 1 \times 5 + 0 \times 2$
$\dot{v}_3$	$= \dot{v}_0 \times \cos v_0$	$= 0 \times \cos 5$
$\dot{v}_4?$		

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace			Forward Tangent (Derivative) Trace		
v_{-1}	$= x_1$	$= 2$	$\dot{v}_{-1}$	$= \dot{x}_1$	$= 1$
v_0	$= x_2$	$= 5$	$\dot{v}_0$	$= \dot{x}_2$	$= 0$
v_1	$= \ln v_{-1}$	$= \ln 2$	$\dot{v}_1$	$= \dot{v}_{-1}/v_{-1}$	$= 1/2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$	$\dot{v}_2$	$= \dot{v}_{-1} \times v_0 + \dot{v}_0 \times v_{-1}$	$= 1 \times 5 + 0 \times 2$
v_3	$= \sin v_0$	$= \sin 5$	$\dot{v}_3$	$= \dot{v}_0 \times \cos v_0$	$= 0 \times \cos 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$	$\dot{v}_4$	$= \dot{v}_1 + \dot{v}_2$	$= 0.5 + 5$

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$

Forward Tangent (Derivative) Trace		
$\dot{v}_{-1}$	$= \dot{x}_1$	$= 1$
$\dot{v}_0$	$= \dot{x}_2$	$= 0$
$\dot{v}_1$	$= \dot{v}_{-1}/v_{-1}$	$= 1/2$
$\dot{v}_2$	$= \dot{v}_{-1} \times v_0 + \dot{v}_0 \times v_{-1}$	$= 1 \times 5 + 0 \times 2$
$\dot{v}_3$	$= \dot{v}_0 \times \cos v_0$	$= 0 \times \cos 5$
$\dot{v}_4$	$= \dot{v}_1 + \dot{v}_2$	$= 0.5 + 5$
$\dot{v}_5$?		

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$

Forward Tangent (Derivative) Trace		
$\dot{v}_{-1}$	$= \dot{x}_1$	$= 1$
$\dot{v}_0$	$= \dot{x}_2$	$= 0$
$\dot{v}_1$	$= \dot{v}_{-1}/v_{-1}$	$= 1/2$
$\dot{v}_2$	$= \dot{v}_{-1} \times v_0 + \dot{v}_0 \times v_{-1}$	$= 1 \times 5 + 0 \times 2$
$\dot{v}_3$	$= \dot{v}_0 \times \cos v_0$	$= 0 \times \cos 5$
$\dot{v}_4$	$= \dot{v}_1 + \dot{v}_2$	$= 0.5 + 5$
$\dot{v}_5$	$= \dot{v}_4 - \dot{v}_3$	$= 5.5 - 0$

Exercise: Forward Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\dot{x}_1 = 1 \longrightarrow \frac{\delta y}{\delta x_1}$$

Forward Primal Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$
y	$= v_5$	$= 11.652$

Forward Tangent (Derivative) Trace		
$\dot{v}_{-1}$	$= \dot{x}_1$	$= 1$
$\dot{v}_0$	$= \dot{x}_2$	$= 0$
$\dot{v}_1$	$= \dot{v}_{-1} / v_{-1}$	$= 1/2$
$\dot{v}_2$	$= \dot{v}_{-1} \times v_0 + \dot{v}_0 \times v_{-1}$	$= 1 \times 5 + 0 \times 2$
$\dot{v}_3$	$= \dot{v}_0 \times \cos v_0$	$= 0 \times \cos 5$
$\dot{v}_4$	$= \dot{v}_1 + \dot{v}_2$	$= 0.5 + 5$
$\dot{v}_5$	$= \dot{v}_4 - \dot{v}_3$	$= 5.5 - 0$
$\dot{\boldsymbol{y}}$	$= \dot{\boldsymbol{v}}_5$	$= \mathbf{5.5}$

Forward Mode for ML?

$$f : \mathbb{R} \to \mathbb{R}^m$$

$$\frac{\delta y_i}{\delta x}$$

can be computed in one forward pass!

$$f : \mathbb{R}^n \to \mathbb{R}$$

$$\nabla f = (\frac{\delta y}{\delta x_1}, \ldots, \frac{\delta y}{\delta x_n})$$

needs n forward passes!

Functions in ML

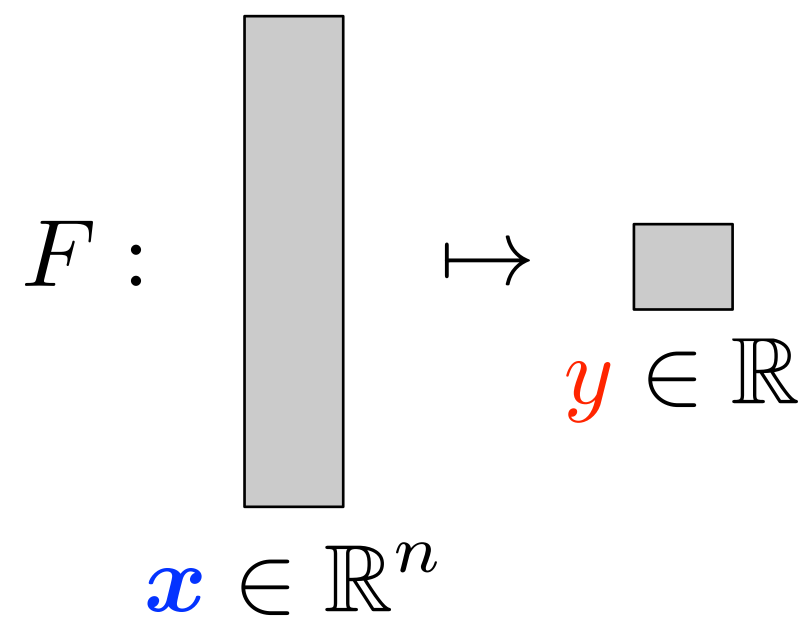
$$f : \mathbb{R}^n \to \mathbb{R}^m$$
$$n \gg m$$

Forward mode AD is not scalable to input dimensionality

Functions in ML

even more extreme, m=1

$$F : \mathbb{R}^n \to \mathbb{R}$$



$$F = D \circ C \circ B \circ A$$

$$y = F(\boldsymbol{x}) = D(C(B(A(\boldsymbol{x}))))$$

$$y = D(\boldsymbol{c}), \quad \boldsymbol{c} = C(\boldsymbol{b}), \quad \boldsymbol{b} = B(\boldsymbol{a}), \quad \boldsymbol{a} = A(\boldsymbol{x})$$

$$y = D(\boldsymbol{c}), \quad \boldsymbol{c} = C(\boldsymbol{b}), \quad \boldsymbol{b} = B(\boldsymbol{a}), \quad \boldsymbol{a} = A(\boldsymbol{x})$$

$$F'(\boldsymbol{x}) = \frac{\partial y}{\partial \boldsymbol{x}} = \left[\frac{\partial y}{\partial x_1} \quad \cdots \quad \frac{\partial y}{\partial x_n} \right]$$

$$F'(\boldsymbol{x}) = \quad \frac{\partial y}{\partial \boldsymbol{c}} \; \frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} \; \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} \; \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}}$$

$$y = D(\boldsymbol{c}), \quad \boldsymbol{c} = C(\boldsymbol{b}), \quad \boldsymbol{b} = B(\boldsymbol{a}), \quad \boldsymbol{a} = A(\boldsymbol{x})$$

$$F'(\boldsymbol{x}) = \frac{\partial y}{\partial \boldsymbol{x}} = \left[\frac{\partial y}{\partial x_1} \quad \cdots \quad \frac{\partial y}{\partial x_n} \right]$$

$$F'(\boldsymbol{x}) = \quad \frac{\partial y}{\partial \boldsymbol{c}} \; \frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} \; \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} \; \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}}$$

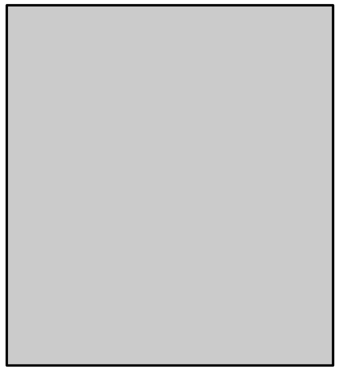
$$\frac{\partial y}{\partial \boldsymbol{c}} = D'(\boldsymbol{c})$$

$$y = D(\boldsymbol{c}), \quad \boldsymbol{c} = C(\boldsymbol{b}), \quad \boldsymbol{b} = B(\boldsymbol{a}), \quad \boldsymbol{a} = A(\boldsymbol{x})$$

$$F'(\boldsymbol{x}) = \frac{\partial y}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \cdots & \frac{\partial y}{\partial x_n} \end{bmatrix}$$

$$F'(\boldsymbol{x}) = \quad \frac{\partial y}{\partial \boldsymbol{c}} \; \frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} \; \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} \; \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}}$$

$$\frac{\partial y}{\partial \boldsymbol{c}} = D'(\boldsymbol{c}) \qquad\qquad \frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} = C'(\boldsymbol{b})$$

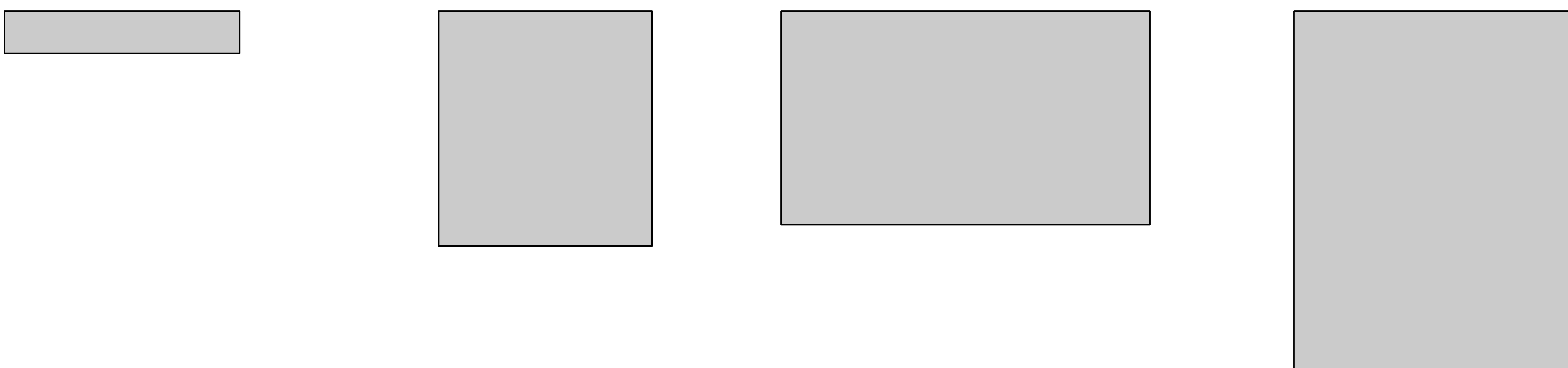


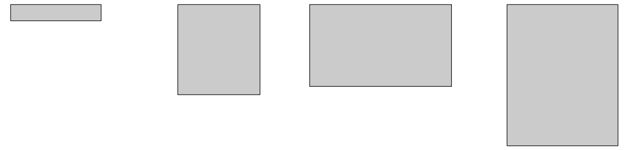
$$y = D(\boldsymbol{c}), \quad \boldsymbol{c} = C(\boldsymbol{b}), \quad \boldsymbol{b} = B(\boldsymbol{a}), \quad \boldsymbol{a} = A(\boldsymbol{x})$$

$$F'(\boldsymbol{x}) = \frac{\partial y}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \cdots & \frac{\partial y}{\partial x_n} \end{bmatrix}$$

$$F'(\boldsymbol{x}) = \quad \frac{\partial y}{\partial \boldsymbol{c}} \; \frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} \; \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} \; \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}}$$

$$\frac{\partial y}{\partial \boldsymbol{c}} = D'(\boldsymbol{c}) \qquad \frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} = C'(\boldsymbol{b}) \qquad \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} = B'(\boldsymbol{a}) \qquad \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}} = A'(\boldsymbol{x})$$

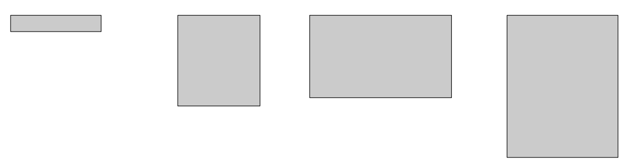




$$F'(\boldsymbol{x}) = \frac{\partial y}{\partial \boldsymbol{c}}\left(\frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}}\left(\underbrace{\frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}}\quad\frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}}}\right)\right)$$

$$\frac{\partial \boldsymbol{b}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial b_1}{\partial x_1} & \cdots & \frac{\partial b_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial b_m}{\partial x_1} & \cdots & \frac{\partial b_m}{\partial x_n} \end{bmatrix}$$

Forward accumulation



$$F'(\boldsymbol{x}) = \left(\left(\underbrace{\frac{\partial y}{\partial \boldsymbol{c}}\quad\frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}}}\right)\frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}}\right)\frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}}$$

$$\frac{\partial y}{\partial \boldsymbol{b}} = \begin{bmatrix} \frac{\partial y}{\partial b_1} & \cdots & \frac{\partial y}{\partial b_m} \end{bmatrix}$$

Reverse accumulation

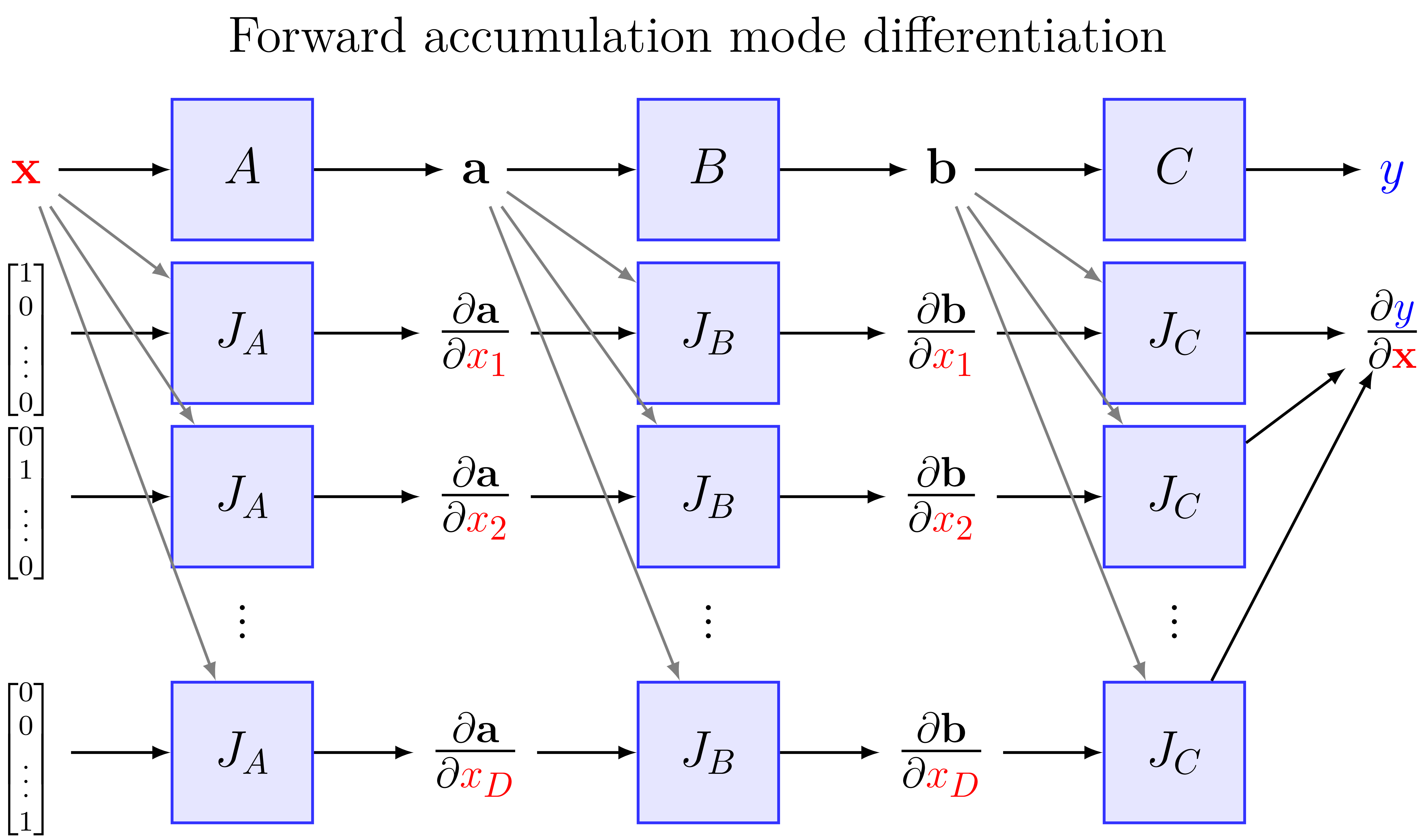
$$F'(\boldsymbol{x})\,\boldsymbol{v} = \quad \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{c}} \; \frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} \; \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} \; \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}} \,\boldsymbol{v}$$

$$F'(\boldsymbol{x})\,\boldsymbol{v} = \quad \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{c}} \left(\frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} \left(\frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} \left(\frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}} \,\boldsymbol{v} \right)\right)\right)$$

Forward accumulation $\leftrightarrow$ Jacobian-vector products

Build Jacobian one column at a time

$$F'(\boldsymbol{x}) \quad = \quad \frac{\partial y}{\partial \boldsymbol{c}} \left(\frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} \left(\frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} \left(\frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}} \; \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x}} \right)\right)\right)$$



Forward accumulation mode differentiation
$\mathbf{x}$
A
$\mathbf{a}$
B
$\mathbf{b}$
C
y
J_A
J_B
J_C
$\frac{\partial \mathbf{a}}{\partial x_1}$
$\frac{\partial \mathbf{b}}{\partial x_1}$
$\frac{\partial \mathbf{a}}{\partial x_2}$
$\frac{\partial \mathbf{b}}{\partial x_2}$
$\frac{\partial \mathbf{a}}{\partial x_D}$
$\frac{\partial \mathbf{b}}{\partial x_D}$
$\frac{\partial y}{\partial \mathbf{x}}$
$\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}$
$\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix}$
$\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}$

$$\boldsymbol{v}^{\mathsf{T}} F'(\boldsymbol{x}) = \boldsymbol{v}^{\mathsf{T}} \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{c}} \quad \frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} \quad \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} \quad \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}}$$

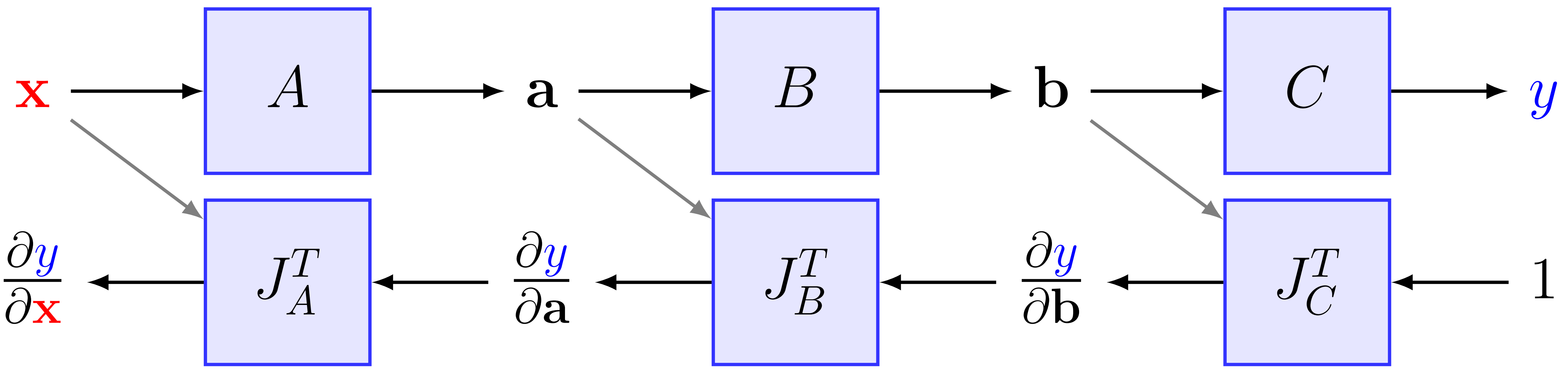
$$\boldsymbol{v}^{\mathsf{T}} F'(\boldsymbol{x}) = \left(\left(\left(\boldsymbol{v}^{\mathsf{T}} \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{c}} \right) \frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} \right) \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} \right) \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}} \Bigg) \Bigg) \Bigg)$$

Reverse accumulation $\leftrightarrow$ vector-Jacobian products

Build Jacobian one row at a time

$$F'(\boldsymbol{x}) = \left(\left(\left(\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{y}} \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{c}} \right) \frac{\partial \boldsymbol{c}}{\partial \boldsymbol{b}} \right) \frac{\partial \boldsymbol{b}}{\partial \boldsymbol{a}} \right) \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}} \Bigg) \Bigg) \Bigg)$$

Reverse accumulation mode differentiation



Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \frac{\delta y}{\delta x_1}, \frac{\delta y}{\delta x_2}$$

both in one reverse pass!

Forward Evaluation Trace			Reverse Adjoint Trace
v_{-1}	$= x_1$	$= 2$	
v_0	$= x_2$	$= 5$	
v_1	$= \ln v_{-1}$	$= \ln 2$	
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$	
v_3	$= \sin v_0$	$= \sin 5$	
v_4	$= v_1 + v_2$	$= 0.693 + 10$	
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$	
y	$= v_5$	$= 11.652$	

Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \frac{\delta y}{\delta x_1}, \frac{\delta y}{\delta x_2}$$

both in one reverse pass!

Forward Evaluation Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$
y	$= v_5$	$= 11.652$

Reverse Adjoint Trace			
$\bar{v}_5$	$= \bar{y}$	$= 1$	$(\partial y / \partial y)$

Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \frac{\delta y}{\delta x_1}, \frac{\delta y}{\delta x_2}$$

both in one reverse pass!

Forward Evaluation Trace

v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$
y	$= v_5$	$= 11.652$

Reverse Adjoint Trace

$\bar{v}_3 ?$

$\bar{v}_4$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_4}$	$= \bar{v}_5 \times 1$	$= 1$
$\bar{v}_5$	$= \bar{y}$	$= 1$	

Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \frac{\delta y}{\delta x_1}, \frac{\delta y}{\delta x_2}$$

both in one reverse pass!

Forward Evaluation Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$
y	$= v_5$	$= 11.652$

Reverse Adjoint Trace			
$\bar{v}_1$?			
$\bar{v}_3$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_3}$	$= \bar{v}_5 \times (-1)$	$= -1$
$\bar{v}_4$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_4}$	$= \bar{v}_5 \times 1$	$= 1$
$\bar{v}_5$	$= \bar{y}$	$= 1$	

Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \frac{\delta y}{\delta x_1}, \frac{\delta y}{\delta x_2}$$

both in one reverse pass!

Forward Evaluation Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$
y	$= v_5$	$= 11.652$

Reverse Adjoint Trace			
$\bar{v}_2?$			
$\bar{v}_1$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_1}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_3$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_3}$	$= \bar{v}_5 \times (-1)$	$= -1$
$\bar{v}_4$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_4}$	$= \bar{v}_5 \times 1$	$= 1$
$\bar{v}_5$	$= \bar{y}$	$= 1$	

Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \frac{\delta y}{\delta x_1}, \frac{\delta y}{\delta x_2}$$

both in one reverse pass!

Forward Evaluation Trace		
$v_{-1} = x_1$		$= 2$
$v_0 = x_2$		$= 5$
$v_1 = \ln v_{-1}$		$= \ln 2$
$v_2 = v_{-1} \times v_0$		$= 2 \times 5$
$v_3 = \sin v_0$		$= \sin 5$
$v_4 = v_1 + v_2$		$= 0.693 + 10$
$v_5 = v_4 - v_3$		$= 10.693 + 0.959$
$y = v_5$		$= 11.652$

Reverse Adjoint Trace		
$\bar{v}_0?$		
$\bar{v}_2 = \bar{v}_4 \frac{\partial v_4}{\partial v_2}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_1 = \bar{v}_4 \frac{\partial v_4}{\partial v_1}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_3 = \bar{v}_5 \frac{\partial v_5}{\partial v_3}$	$= \bar{v}_5 \times (-1)$	$= -1$
$\bar{v}_4 = \bar{v}_5 \frac{\partial v_5}{\partial v_4}$	$= \bar{v}_5 \times 1$	$= 1$
$\bar{v}_5 = \bar{y}$	$= 1$	

Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \frac{\delta y}{\delta x_1}, \frac{\delta y}{\delta x_2}$$

both in one reverse pass!

Forward Evaluation Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
$\boxed{v_3 = \sin v_0}$		$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$
y	$= v_5$	$= 11.652$

Reverse Adjoint Trace

$\bar{v}_{-1}$?

$\bar{v}_0$	$= \bar{v}_3 \frac{\partial v_3}{\partial v_0}$	$= \boxed{\bar{v}_3 \times \cos v_0}$	$= -0.284$
$\bar{v}_2$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_2}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_1$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_1}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_3$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_3}$	$= \bar{v}_5 \times (-1)$	$= -1$
$\bar{v}_4$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_4}$	$= \bar{v}_5 \times 1$	$= 1$
$\bar{v}_5$	$= \bar{y}$	$= 1$	

Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \frac{\delta y}{\delta x_1}, \frac{\delta y}{\delta x_2}$$

both in one reverse pass!

Forward Evaluation Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
$\boxed{v_2}$	$\boxed{= v_{-1} \times v_0}$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$
y	$= v_5$	$= 11.652$

Reverse Adjoint Trace			
$\bar{v}_0 ?$			
$\bar{v}_{-1}$	$= \bar{v}_2 \frac{\partial v_2}{\partial v_{-1}}$	$= \boxed{\bar{v}_2 \times v_0}$	$= 5$
$\bar{v}_0$	$= \bar{v}_3 \frac{\partial v_3}{\partial v_0}$	$= \bar{v}_3 \times \cos v_0$	$= -0.284$
$\bar{v}_2$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_2}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_1$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_1}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_3$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_3}$	$= \bar{v}_5 \times (-1)$	$= -1$
$\bar{v}_4$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_4}$	$= \bar{v}_5 \times 1$	$= 1$
$\bar{v}_5$	$= \bar{y}$	$= 1$	

Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \frac{\delta y}{\delta x_1}, \frac{\delta y}{\delta x_2}$$

both in one reverse pass!

Forward Evaluation Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= \boxed{v_{-1} \times v_0}$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$
y	$= v_5$	$= 11.652$

Reverse Adjoint Trace

$\bar{v}_{-1}$?

$\bar{v}_0$	$= \bar{v}_0 + \bar{v}_2 \frac{\partial v_2}{\partial v_0}$	$= \bar{v}_0 + \boxed{\bar{v}_2 \times v_{-1}}$	$= 1.716$
$\bar{v}_{-1}$	$= \bar{v}_2 \frac{\partial v_2}{\partial v_{-1}}$	$= \bar{v}_2 \times v_0$	$= 5$
$\bar{v}_0$	$= \bar{v}_3 \frac{\partial v_3}{\partial v_0}$	$= \bar{v}_3 \times \cos v_0$	$= -0.284$
$\bar{v}_2$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_2}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_1$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_1}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_3$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_3}$	$= \bar{v}_5 \times (-1)$	$= -1$
$\bar{v}_4$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_4}$	$= \bar{v}_5 \times 1$	$= 1$
$\bar{v}_5$	$= \bar{y}$	$= 1$	

Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \frac{\delta y}{\delta x_1}, \frac{\delta y}{\delta x_2}$$

both in one reverse pass!

Forward Evaluation Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
$\boxed{v_1 = \ln v_{-1}}$		$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$
y	$= v_5$	$= 11.652$

Reverse Adjoint Trace			
$\bar{v}_{-1}$	$= \bar{v}_{-1} + \bar{v}_1 \frac{\partial v_1}{\partial v_{-1}}$	$= \bar{v}_{-1} + \boxed{\bar{v}_1 / v_{-1}}$	$= 5.5$
$\bar{v}_0$	$= \bar{v}_0 + \bar{v}_2 \frac{\partial v_2}{\partial v_0}$	$= \bar{v}_0 + \bar{v}_2 \times v_{-1}$	$= 1.716$
$\bar{v}_{-1}$	$= \bar{v}_2 \frac{\partial v_2}{\partial v_{-1}}$	$= \bar{v}_2 \times v_0$	$= 5$
$\bar{v}_0$	$= \bar{v}_3 \frac{\partial v_3}{\partial v_0}$	$= \bar{v}_3 \times \cos v_0$	$= -0.284$
$\bar{v}_2$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_2}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_1$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_1}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_3$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_3}$	$= \bar{v}_5 \times (-1)$	$= -1$
$\bar{v}_4$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_4}$	$= \bar{v}_5 \times 1$	$= 1$
$\bar{v}_5$	$= \bar{y}$	$= 1$	

Exercise: Reverse Mode

$$y = f(x_1, x_2) = ln(x_1) + x_1 x_2 - sin(x_2)$$

Solve at point $(x_1, x_2) = (2, 5)$

$$\bar{y} = \frac{\delta y}{\delta y} = 1 \longrightarrow \underset{(\bar{x}_1)}{\frac{\delta y}{\delta x_1}}, \underset{(\bar{x}_2)}{\frac{\delta y}{\delta x_2}}$$

both in one reverse pass!

Forward Evaluation Trace		
v_{-1}	$= x_1$	$= 2$
v_0	$= x_2$	$= 5$
v_1	$= \ln v_{-1}$	$= \ln 2$
v_2	$= v_{-1} \times v_0$	$= 2 \times 5$
v_3	$= \sin v_0$	$= \sin 5$
v_4	$= v_1 + v_2$	$= 0.693 + 10$
v_5	$= v_4 - v_3$	$= 10.693 + 0.959$
y	$= v_5$	$= 11.652$

Reverse Adjoint Trace			
$\boldsymbol{\bar{x}_1}$	$\boldsymbol{= \bar{v}_{-1}}$		$\boldsymbol{= 5.5}$
$\boldsymbol{\bar{x}_2}$	$\boldsymbol{= \bar{v}_0}$		$\boldsymbol{= 1.716}$
$\bar{v}_{-1}$	$= \bar{v}_{-1} + \bar{v}_1 \frac{\partial v_1}{\partial v_{-1}}$	$= \bar{v}_{-1} + \bar{v}_1 / v_{-1}$	$= 5.5$
$\bar{v}_0$	$= \bar{v}_0 + \bar{v}_2 \frac{\partial v_2}{\partial v_0}$	$= \bar{v}_0 + \bar{v}_2 \times v_{-1}$	$= 1.716$
$\bar{v}_{-1}$	$= \bar{v}_2 \frac{\partial v_2}{\partial v_{-1}}$	$= \bar{v}_2 \times v_0$	$= 5$
$\bar{v}_0$	$= \bar{v}_3 \frac{\partial v_3}{\partial v_0}$	$= \bar{v}_3 \times \cos v_0$	$= -0.284$
$\bar{v}_2$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_2}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_1$	$= \bar{v}_4 \frac{\partial v_4}{\partial v_1}$	$= \bar{v}_4 \times 1$	$= 1$
$\bar{v}_3$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_3}$	$= \bar{v}_5 \times (-1)$	$= -1$
$\bar{v}_4$	$= \bar{v}_5 \frac{\partial v_5}{\partial v_4}$	$= \bar{v}_5 \times 1$	$= 1$
$\bar{v}_5$	$= \bar{y}$	$= 1$	

Backpropagation is a special case of Reverse Mode AD