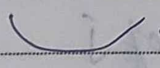
Gradient Descent.

 → convex f^n  → non-convex

Refer ML notes — repetitive.

Derivatives

→ explained how ~~et f^n~~ y changes with x
and how we track the change using slope.

More Derivative examples.

ehhhh..........

waid.

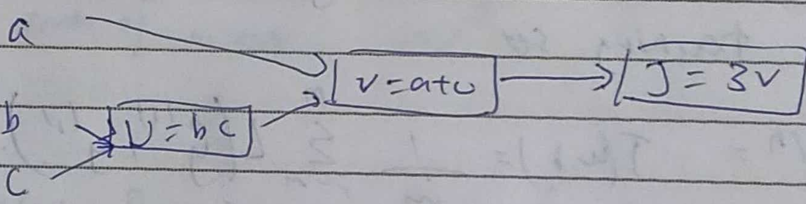
Computation Graph.

computing a f^n consists of multiple steps
lets say

$$J(a,b,c) = \underbrace{3\underbrace{(a+\underbrace{bc}_{u})}_{v}}_{J}$$

$u = bc$
$v = a+u$
$J = 3v$

computation graph of above



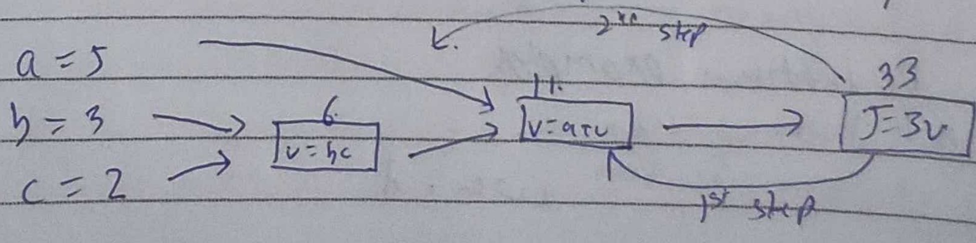
this graph is useful when you want to ~~of~~ OPTIMIZE one of the output variable say v or J

when we go from left to right → we can calculate value of J

To compute derivaties → we go from right to left

Derivatives with a computation Graph

* One step of backward propagation on a computation graph yields derivative of a final output variable



1st step → we want to know how does J behave when we tweak values of v.

$J = 3v$ | $f(a) = 3a$

$v = 11 \rightarrow 11.001$ | $\frac{df(a)}{da} = 3$ $\therefore \frac{dJ}{dv} = ? = 3$

$J = 33 \rightarrow 33.003$

i.e. when we change values of v, J changes itself 3 times those values.

2nd step → now, how does J behave, when we change values of a?

$a = 5 \rightarrow 5.001$

$v = 11 \rightarrow 11.001$ ⟶ we conclude J changes 3 times informally: $\frac{dJ}{da} = 3$

$J = 33 \rightarrow 33.003$

gist is → changing a → changes v by some amount → changes J 3 times

$\frac{dv}{da} = 1$

~~formally~~:

$$\frac{dJ}{da} = \underbrace{\frac{dJ}{dv} \cdot \frac{dv}{da}}_{1 \times 3 = 3} \longrightarrow \text{chain rule!!}$$

This is how we ~~change~~ understand how changing previous variables, ~~to~~ changes the value of final output variable

* In NNs, final output variable is our output and we try to understand how changing ~~[illegible]~~ previous layer parameter changes our output

DOMS

3rd example

how does J behave when we change b?

changing b → changes u → changes v → changes v
↓
changes J

so $\frac{\partial J}{\partial d} = \frac{\partial J}{\partial v} \frac{\partial v}{\partial v} \frac{\partial u}{\partial b} = \cancel{x} 3 \times 2 = 6$

similarly

$$\frac{\partial J}{\partial c} = \frac{\partial J}{\partial u} \cdot \frac{\partial u}{\partial c} = 9$$