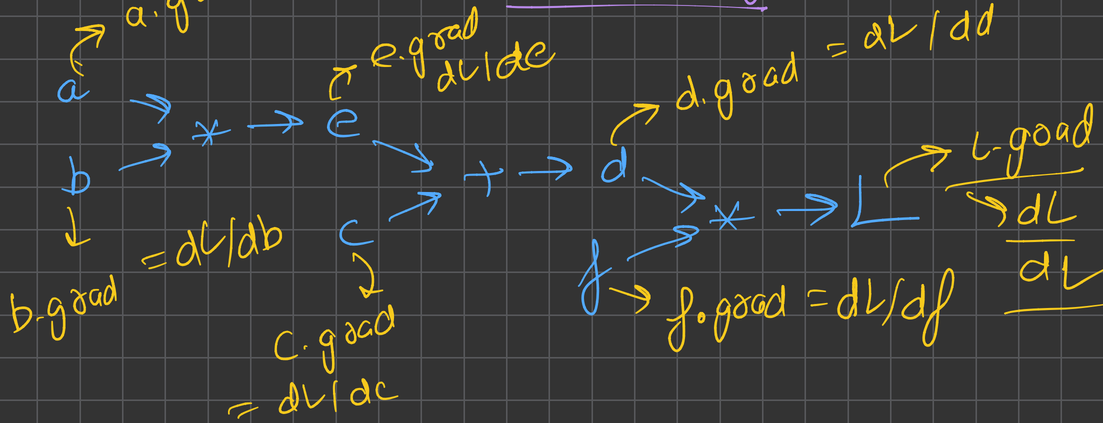
Backward Propagation



---

why we set grad = 0 initially?

when slope of a $f(x)$ w.r.t. x is 0, at a certain pt., it means on changing x by tiny amt $(x \to x+h)$, $(x \to x-h)$, $f(x)$ remain same. Eg:- $dL/da = 0$, L is not changing on changing a by h, $(a+h)$ $(a-h)$.

A) L.grad = ? $\quad dL/dL$

$$\text{slope} = \frac{f(x+h) - f(x)}{h} = \frac{(\cancel{L}+h) - \cancel{L}}{h} = 1$$

$\boxed{\text{slope} = 1}$ → Analytical

B) $\frac{dL}{dL} = \frac{dL}{dL} = 1$ (How does L changes w.r.t. L)

$\boxed{\text{slope} = 1}$ → calculus

Made with Goodnotes

2) d.grad = ? $dL/dd = ?$, $L = d * f$

A) $\text{slope} = \frac{f(x+h) - f(x)}{h} = \frac{(d+h)*f - d*f}{h}$

$= \frac{\cancel{d*f} + \cancel{h}*f - \cancel{d*f}}{\cancel{h}} = \boxed{f}$

$\boxed{\text{slope} = f}$ → Analytical

B) $dL/dd = \frac{d(d*f)}{dd} = d\cancel{\frac{df}{dd}}^{0} + f\cancel{\frac{dd}{dd}}^{1}$ $(f = x = -2)$

$\boxed{\text{slope} = f}$ → calculus

$\frac{dL}{df}$ = f.grad = ? $L = d * f$

$\frac{dL}{df} = \text{slope} = \frac{f(x+h) - f(x)}{h} = \frac{d(f+h) - df}{h}$

$= \frac{\cancel{df} + d\cancel{h} - \cancel{df}}{\cancel{h}} = d$, $\boxed{\text{slope} = d}$

→ Analytically

B) $\frac{dL}{df} = \frac{d}{df}(d*f) = d\cancel{\frac{df}{df}}^{1} + f\cancel{\left(\frac{dd}{df}\right)}^{0}$ $(d = x = 4)$

$= d$

$\boxed{\text{slope} = d}$ → calculus