 2/22/2022

Path weights on a graph

Forward

$$(start, \dots) \quad \begin{matrix} (arg_1, \frac{darg_1}{dstart}) \xrightarrow{\partial f / \partial arg_1} \\ (arg_2, darg_2/dstart) \longrightarrow \\ (arg_k, \frac{darg_k}{dstart}) \xrightarrow{\partial f / \partial arg_k} \end{matrix} \quad (f(arg_1, \dots, arg_k), ?) \quad \dots \; end$$

$$\frac{df}{dend} = \frac{\partial f}{\partial arg_1}\frac{darg_1}{dstart} + \dots + \frac{\partial f}{\partial arg_k}\frac{darg_k}{dstart}$$

Scalars: all quantities can be scalar

Vectors: $(start, vector) \xrightarrow{\text{Jacobians}}$ directional derivative

$(start, J) \longrightarrow$ Jacobian

Anything: can be operators

Matrix (vector) + ... + Matrix (vector)

One directional derivative

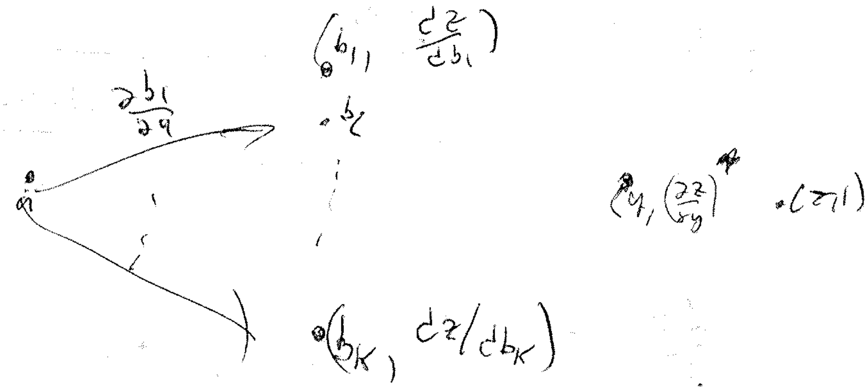
Matrix (Matrix) + ... + Matrix (Matrix)

Jacobian

Forward to get gradient of loss

$$(start, J) \xrightarrow{\text{lots of Jacobians}} (end\text{-}1, Matrix) \xrightarrow{(n \times 1)} gradient^T$$

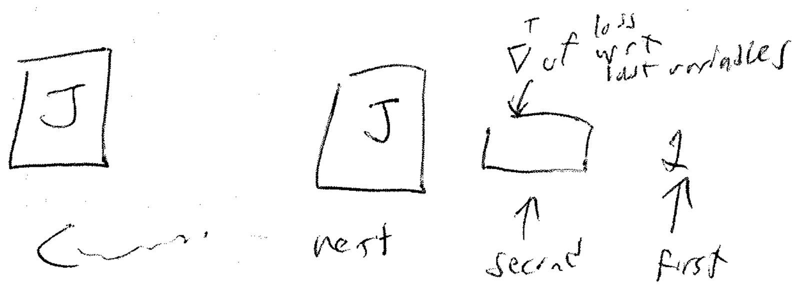
Reverse Mode



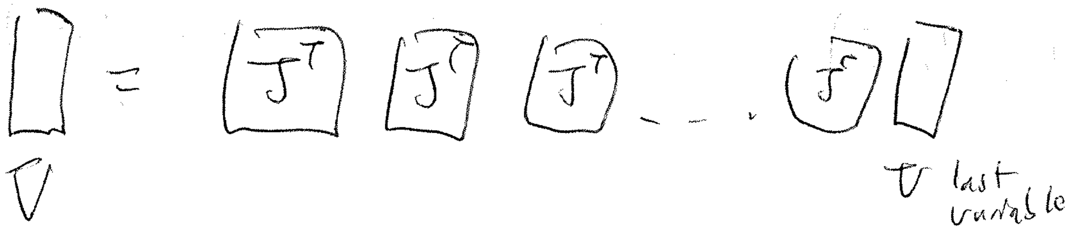
All scalars

$$\frac{dz}{da} = \frac{\partial b_1}{\partial a}\frac{dz}{db_1} + \cdots + \frac{\partial b_K}{\partial a}\frac{dz}{db_K}$$

Reverse Mode looks like this for scalar fnct of vectors



Many prefer transposing



Constraint: Executions in Software

Forward: Run through Program

"Synchronous"

do all three at the same time as part of $f(\text{arg}_1, \ldots, \text{arg}_k)$

$$\sum \frac{\partial f}{\partial \text{arg}_i} \frac{d\,\text{arg}_i}{d\,\text{start}}$$ the obvious way

Reverse

Initially run through program in forward mode build the dag with the labels

initialize all $\frac{d\,\text{end}}{d\,\text{vars}} \to 0$ $\quad \frac{dz}{dz} = 1$ (empty path)

As you traverse backwards you asynchronousl update

e.g. $$\frac{dz}{da} \mathrel{+}= \frac{\partial b_j}{\partial a} \frac{dz}{db_j}$$

Example

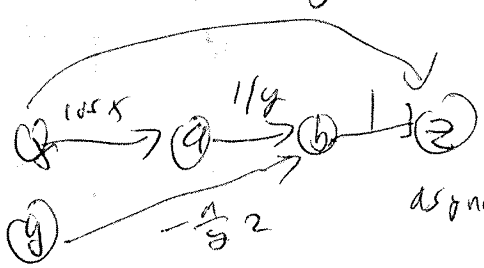
input x, y

(1) $a = \sin x$

(2) $b = a/y$

(3) $z = b + x$

return z



asynchronous
$(z, 1)$
$(x, 1)$
$(b, 1)$
$(a, 1/y)$
$(x, 1 + \frac{\cos x}{y})$

vs
$(a, \cos x)$
$(b, \cos x / y)$
$(z, \frac{\cos x}{y} + 1) \leftarrow$ synchonous

$+, -, \div, \sqrt{\ }$, reverse

$$
\begin{tikzcd}
(a, p_1) \arrow[dr, "1"] & \\
& (a \pm b, q) \\
(b, p_2) \arrow[ur, "-1"'] &
\end{tikzcd}
$$

$p_1 \mathrel{+}= q$

$p_2 \mathrel{\mp}= q$

$$
\begin{tikzcd}
(a, p_1) \arrow[dr, "b"] & \\
& (a * b, q) \\
(b, p_2) \arrow[ur, "a"'] &
\end{tikzcd}
$$

$p_1 \mathrel{+}= b\, q$

$p_2 \mathrel{+}= a\, q$

$$
\begin{tikzcd}
(a, p_1) \arrow[dr, "1/b"] & \\
& (a/b, q) \\
(b, p_2) \arrow[ur, "-\frac{a}{b^2}"'] &
\end{tikzcd}
$$

$p_1 \mathrel{\cancel{*}{+}}= \frac{1}{b} \cdot q$

$p_2 \mathrel{+}= \frac{-a}{b^2}\, q$

scalar functions e.g.

$$(a, p) \xrightarrow{\cos a} (\sin a, q)$$

$p \mathrel{+}= (\cos a)\, q$

(3) 2/22/2022

PINN2

Where is the "informed" in PINN?

Hooke's Law Example

We suspect $u'' \approx ku$ but not exactly

e.g. $u'' = ku + .01 \sin t$

We'll scrape some data deliberately
chosen so that being non-informed can
never work but being informed does well