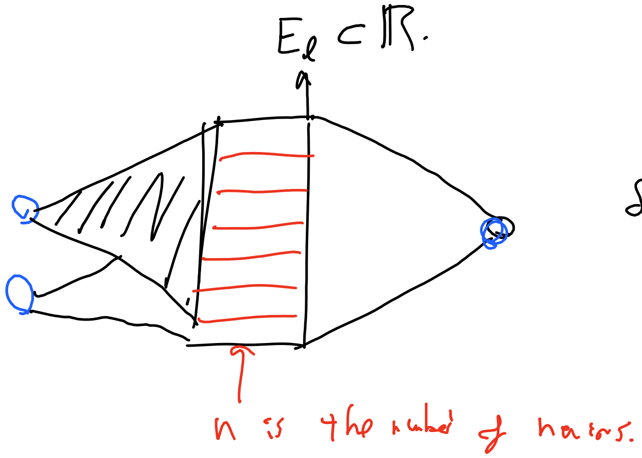
Simple Derivation For Single Layers

Saturday, June 24, 2017 2:54 PM

Let $D : \boxed{\mathbb{R}^n} \xrightarrow{\delta} \boxed{C(\mathbb{R})} \xrightarrow{f} \boxed{\mathbb{R}^m}$ be a deep func

Then let N be the n-discrete instantiation of D.



$$f : \xi \longmapsto \int \xi \cdot w^1_t \, d\mu(u)$$

$$\delta : x \longmapsto \sum_{j=1}^{|x|} x_j w^\ell_j(u)$$

Let $w^0_j = \sum_{n=1}^{\infty} \chi_{nI}(u) w_{nj}$,

Therefore $\sigma \circ \delta(x) = \sigma \circ \sum_{j=1} x_j \sum_{n=1}^{\infty} \chi_{nI}(u) w_{nj}$

$$= \sigma \circ \sum_{n=1}^{\infty} (W^0 x)_n \chi_{nI}(u) . \longleftarrow u \in E_\ell$$

Let $w^1_0 = \sum \chi_{nI}(u) w^1_{n,1}$, $(w_n) \in \mathbb{R}^{|N|} \otimes \mathbb{R}$.

Therefore

$$h \circ \sigma \circ \delta = \int_{E_\ell} \sum_{n=1} \chi_{nI}(u) w^1_n \sum_{m=1}^{\infty} (W^0 x)_n \chi_{mI}(u) \, d\mu$$

$$= \int_{E_\ell} \sum_{n=1} \chi_{nI}(u) w^1_n \sigma((W^0 x)_n) \, d\mu(n)$$

$$= \sum w^1_n \sigma((W^0 x)_n) = W^1 \sigma(W^0 x).$$

$\uparrow$ $0 \wedge \mathbb{R}^{\infty}$ $\quad$ $\mathbb{R}^{\infty} \times I$

Now we derive a continuous parameterization of t weights by changing the integral of integration.

Suppose $W^2 \in \mathbb{R}^\infty \otimes \mathbb{R}^{|I|}$; then, $\partial_\ell [D]$ is computed

$$D(x; W^1, W^2, \ell) = \int_0^\ell \sum_{n=1}^\infty \chi_{nI}(u)\, W^1_n \cdot g^1(u)\, d\mu(u).$$

$$\partial_\ell \mathcal{L}(D(x; W^1, W^2, \ell)) = \frac{\partial \mathcal{L}}{\partial y^2} \frac{\partial y^2}{\partial \ell}.$$

$$\uparrow$$

Thus
$$\frac{\partial y^2}{\partial \ell} = \frac{\partial}{\partial \ell} \int_0^\ell \sum_{n=1}^\infty \chi_{nI}\, W^1_n \cdot g^1(u)\, d\mu(u).$$

$$= \sum_{n=1}^\infty \chi_{nI}(\ell)\, W^1_n\, g^1(q)\, d\mu(q).$$

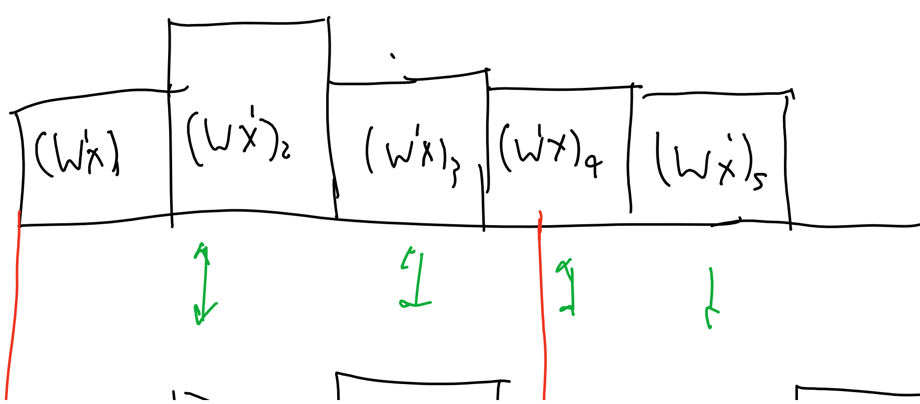
$$= W^1_k\, g^1(\ell) \qquad ; k =$$

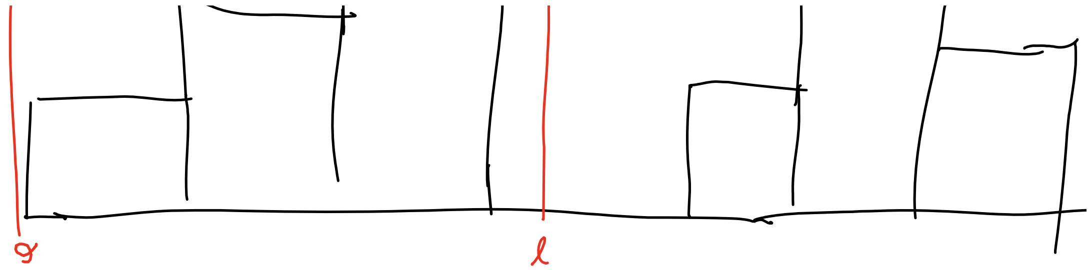
$$= W^1_k\, \sigma\left(\sum_{i=1}^{|I|} W_{ki}\, x_i\right)$$

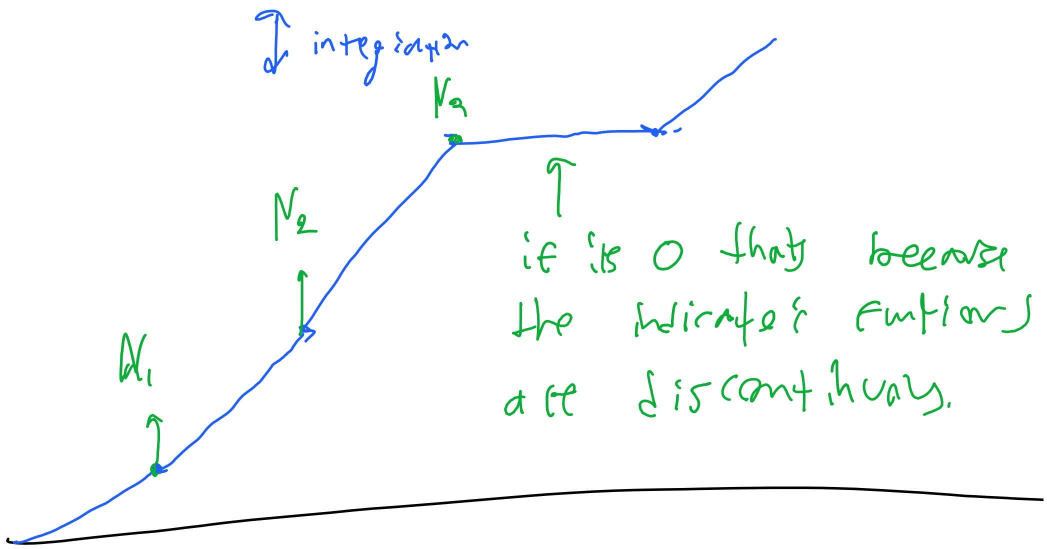
What

Okay so:

$J(x)$

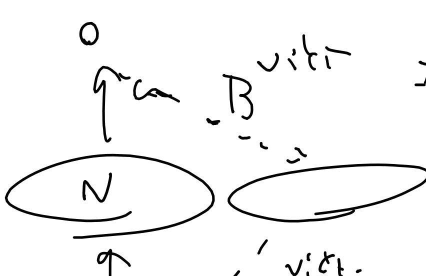




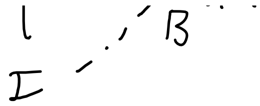


Using the intuition, we want to assume forms where a gradient still exists direction.

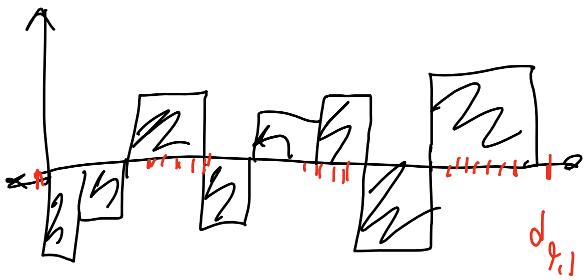
$\boxed{\text{"Neural Tensions"}}$ How do you capture N's fit in a continuous way.



If $\mathcal{L}(N)$ is in a local min virtual gradient ⟹ is m than normal potentially



"Delocalized Interval Shift"



$d_{1,1}, d_{2,}, \ldots \ \partial_E = \partial_E \int_{\mathbb{R}} E \to [0,1] \quad F(u) \Rightarrow$

Descend on all such points which are dense in the measure μ; that is,

$$\partial E = \int_{\mathbb{R}} F(u)\, \partial_E \mu(u) = \int_{\mathbb{R}} f(u) \frac{dE}{du}\, du.$$

· Standard Lebesgue measure.

$\partial E \in f: \mathbb{R} \to [0,1]^*.$

Or assume rest of line is randomly distributed, or is such that the vertical is maximized.