

Lecture 11 slide 25

under the same assumptions of a continuous RW and continuous input signal warped with operator  $\tau(t)$

We will apply Taylor expansion:

$$h(\tau(t+\delta)) \approx h(\tau(t)) + \delta \frac{dh(\tau(t))}{d\tau}$$

Set  $\delta=1$  and solve for  $\frac{dh}{d\tau}$

$$\rightarrow \frac{dh(\tau(t))}{d\tau} = \frac{h(\tau(t+1)) - h(\tau(t))}{1} = R(\tau(t+1), h(\tau(t))) - h(\tau(t)) \rightarrow \textcircled{1}$$

But we want  $\frac{dh}{dt}$ , using the chain rule we know that

$$\frac{dh(\tau(t))}{dt} = \frac{dh(\tau(t))}{d\tau} \frac{d\tau}{dt} \rightarrow \textcircled{2}$$

plug  $\textcircled{1}$  in  $\textcircled{2}$

$$\frac{dh(\tau(t))}{dt} = \frac{d\tau}{dt} R(\tau(t+1), h(\tau(t))) - \frac{d\tau}{dt} h(\tau(t))$$

apply Taylor expansion with  $\delta=1$

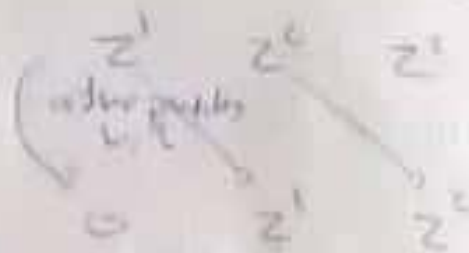
$$h(\tau(t+1)) - h(\tau(t)) = \frac{d\tau}{dt} R(\tau(t+1), h(\tau(t))) - \frac{d\tau}{dt} h(\tau(t))$$

solve for  $h(\tau(t+1))$

$$\rightarrow h(\tau(t+1)) = \frac{d\tau}{dt} R(\tau(t+1), h(\tau(t))) + \left(1 - \frac{d\tau}{dt}\right) h(\tau(t))$$

□

$$\bar{z}(t) = \begin{cases} 0, & t \leq t' \\ z^{(t-t')}, & t > t' \end{cases}$$



let  $\bar{z}^{(0)} = \bar{z}^{(t+1)}$ , we want to show that 1/ ① then

$$h^{(0)} = h^{(t+1)}$$

$$h^{(0)} = R(\bar{z}^{(0)}, h^{(0)}) = R(\bar{z}^{(2)}, h^{(0)}) \rightarrow \textcircled{*}$$

$$h^{(0)} = R(\bar{z}^{(0)}, h^{(0)}) = \text{[scribbled out]}$$

$$= R(\bar{z}^{(0)}, R(\bar{z}^{(0)}, h^{(0)})) \rightarrow \textcircled{**}$$

from ① & ② we want to show that

$$h^{(0)} = R(\bar{z}^{(0)}, h^{(0)}), \text{ where } \bar{z}^{(0)} = \begin{cases} 0 & 1 \leq t' \\ z^{(t-t')} & 1 > t' \end{cases}$$

\* Since the sequence is padded by zeros from the left end the sequence always begins at  $t=1$   $\bar{z}^{(0)} = 0$

$$\Rightarrow h^{(0)} = R(0, h^{(0)})$$

in order for this condition to be satisfied,  $h^{(0)}$  must be constant w.r.t  $R$  and  $R$  will be linear

$$h^{(0)} = W\bar{z}^{(0)} + \mathbf{1}h^{(0)} = h^{(0)}$$

$$\text{and } \bar{z}^{(0)} = 0$$

□

$\Rightarrow$  with padded sequences, RNNs are translation equivariant.

gauge equivariant convolution: (over scalar fields)

$$(K * x)(\vec{r}) = \int_{\mathbb{R}^S} K(v) x(\exp_u(w_u(v))) dv$$

gauge equivariance implies the following:

$$(K * g x)(\vec{r}) = \rho(g) (K * x)$$

$K$  is a function  $\mathbb{R}^S \rightarrow \mathbb{R}$ , the function depends on the coordinates that is defined by the gauge, hence when we apply a gauge transform  $K$  value will change, but  $K$  outputs a scalar value ( $K$  is a scalar), and a scalar field is independent of the gauge transform.

$$K(gv) = K(v)$$

hence a scalar convolution is gauge equivariant, if the filter is gauge invariant.