Exercise

Lecture 11 slide 25

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

We will apply Taylor expansion

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

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

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

$$= R(z(\tau(t+1)), h(\tau(t))) - h(\tau(t)) \rightarrow (1)$$

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

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

plug (1) in (2)

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

$\downarrow$ apply Taylor expansion with $\delta = 1$

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

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

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

$\square$

Scanned with CamScanner

12 slide 14

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

$z^1 \quad z^2 \quad z^3$

after padding w.r.t. t'

$0 \quad z^1 \quad z^2$

let $\underbrace{\bar{z}'^{(t)} = \bar{z}^{(t+1)}}_{①}$, we want to show that if ① then $h'^{(t)} = h^{(t+1)}$..

$h'^{(1)} = R(\bar{z}'^{(1)}, h^{(0)}) = R(\bar{z}^{(2)}, h^{(0)}) \longrightarrow (*)$

$h^{(2)} = R(\bar{z}^{(2)}, h^{(1)})$

$= R(\bar{z}^{(2)}, R(\bar{z}^{(1)}, h^{(0)})) \rightarrow (**)$

from $(*)$ & $(**)$ we want to show that

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

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

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

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

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

and $\bar{z}^{(1)} = 0$ $\square$

$\Rightarrow$ with padded sequences, RNNs are translation equivariant gauge equivariant convolution (over scalar fields)

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

gauge equivariance implies the following:

$$(K * g x)(p) = \rho(g)\,(K * x)$$

K is a function $\mathbb{R}^S \to \mathbb{R}$, 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 scalar field is independent of the gauge transformation
$$K(gv) = K(v)$$
hence a scalar convolution is gauge equivariant if the filter is gauge invariant.