RNNs

Recurrent Neural Networks

St. Joseph's University, Bengaluru

ML 2 BDA3321

Computational Graphs

Node

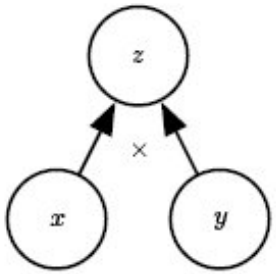
In our graphs, each node will indicate a variable.

Operation

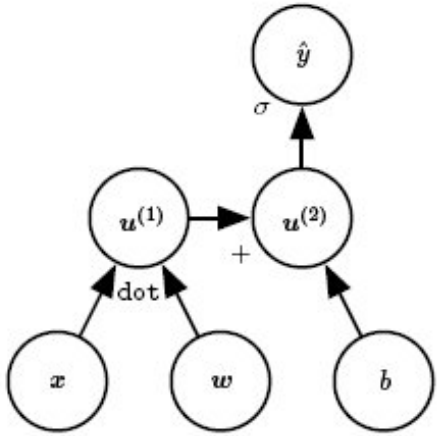
An operation is a function that returns a single value variable.

Computational Graphs

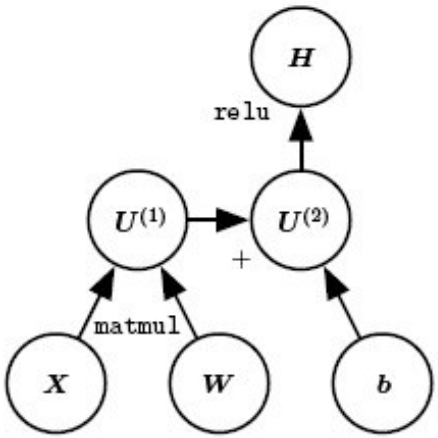
Examples



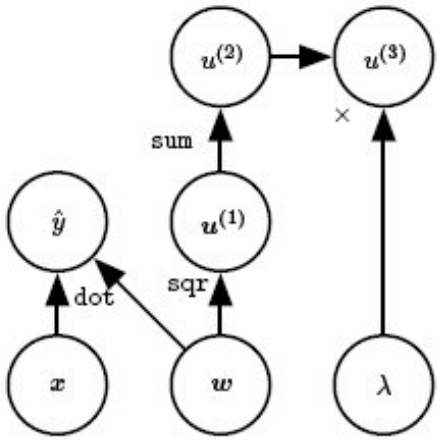
(a)



(b)



(c)



(d)

[Ian17]

Unfolding a recurrent computation

Example 1

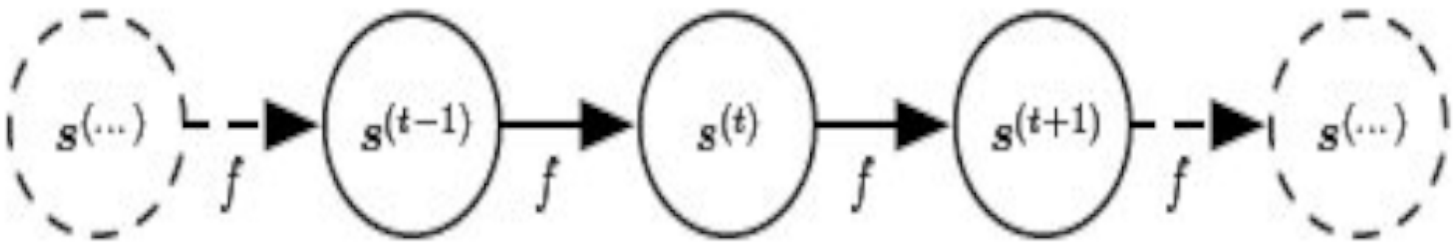
Consider the dynamic system

$$s^{(t)} = f(s^{t-1}; \theta)$$

For three steps, we get:

$$s^{(3)} = f(s^{(2)}; \theta) = f(f(s^{(1)}; \theta))$$

This can be represented graphically as



[Ian17]

Unfolding a recurrent computation

Example 2

Consider a dynamic system driven by an external signal $x^{(t)}$:

$$s^{(t)} = f(s^{(t-1)}, x^{(t)}; \theta)$$

This can be used to define an RNN.

$$h^{(t)} = f(h^{(t-1)}, x^{(t)}; \theta)$$

Representing the unfolded recurrence after t steps with a function $g^{(t)}$:

$$\begin{aligned} h^{(t)} &= g^{(t)}(x^{(t)}, x^{(t-1)}, \ldots, x^{(2)}, x^{(1)}) \\ &= f(h^{(t-1)}, x^{(t)}; \theta) \end{aligned}$$

References I

[Mur12] Kevin P Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.

[Mar14] Stephen Marsland. *Machine Learning, An Algorithmic Perspective*. CRC Press, 2014.

[Ian17] Aaron Courville Ian Goodfellow Yoshua Bengio. *Deep Learning*. MIT Press, 2017.