Introduction to Deep Learning (CS474)

Lecture 22

Rupam

Outline

- **Module 3**
  - **Introduction to Recurrent Neural Network (RNN)**

Rupam

Unfolding Computational Graph

- We refer to RNNs as operating on a sequence that contains vectors $\mathbf{x}^{(t)}$ with the time step index **t** ranging from **1** to **τ**.

- A **computational graph** is a way to formalize the structure of a set of computations, such as those involved in mapping inputs and parameters to outputs and loss.

- For example, consider the classical form of a dynamical system:

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

where $\mathbf{s}^{(t)}$ is called the state of the system.

Equation 1 is recurrent because the definition of **s** at time ***t*** refers back to the same definition at time ***t − 1*** .

Slide credit: I. Goodfellow, Y. Bengio, A. Courville

Rupam

Unfolding Computational Graph

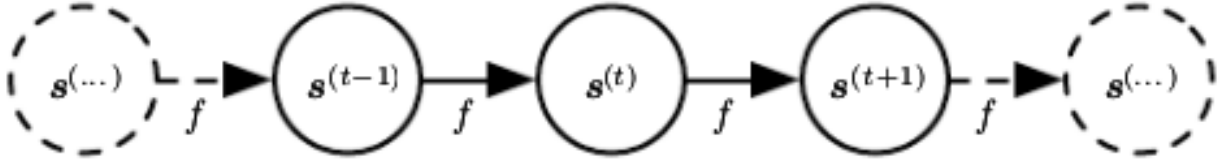


Figure 1: The classical dynamical system described by equation 1, illustrated as an unfolded computational graph. Each node represents the state at some time t and the function f maps the state at t to the state at $t+1$. The same parameters (the same value of $\boldsymbol{\theta}$ used to parametrize f) are used for all time steps.

Slide credit: I. Goodfellow, Y. Bengio, A. Courville

Rupam

Visualizing RNN

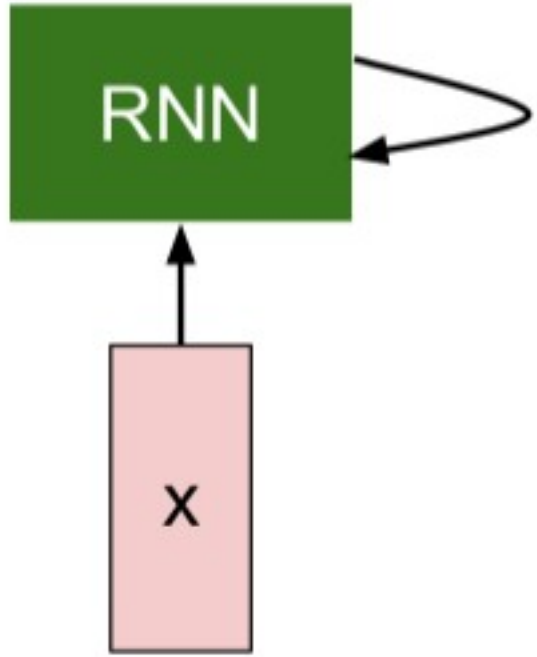


Image credit: F. Li & J. Johnson & S. Yeung

Rupam

Visualizing RNN

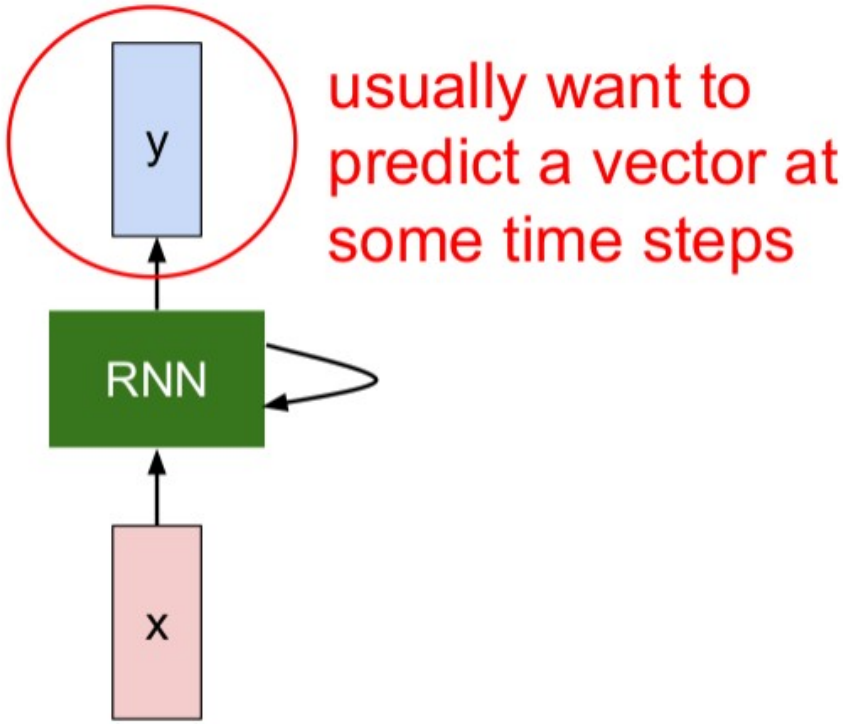


Image credit: F. Li & J. Johnson & S. Yeung

Rupam

Visualizing RNN

We can process a sequence of vectors **x** by applying a **recurrence formula** at every time step:

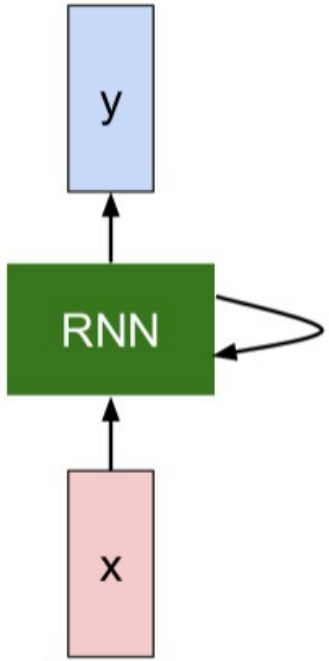
$$h_t = f_W(h_{t-1}, x_t)$$

new state

old state

input vector at some time step

some function with parameters W



Slide credit: F. Li & J. Johnson & S. Yeung

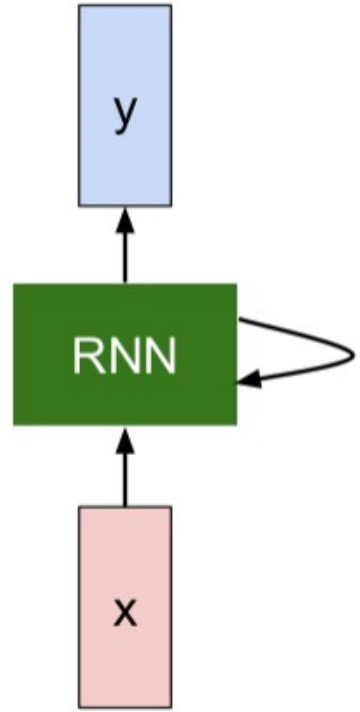
Rupam

Visualizing RNN

We can process a sequence of vectors **x** by applying a **recurrence formula** at every time step:

$$h_t = f_W(h_{t-1}, x_t)$$

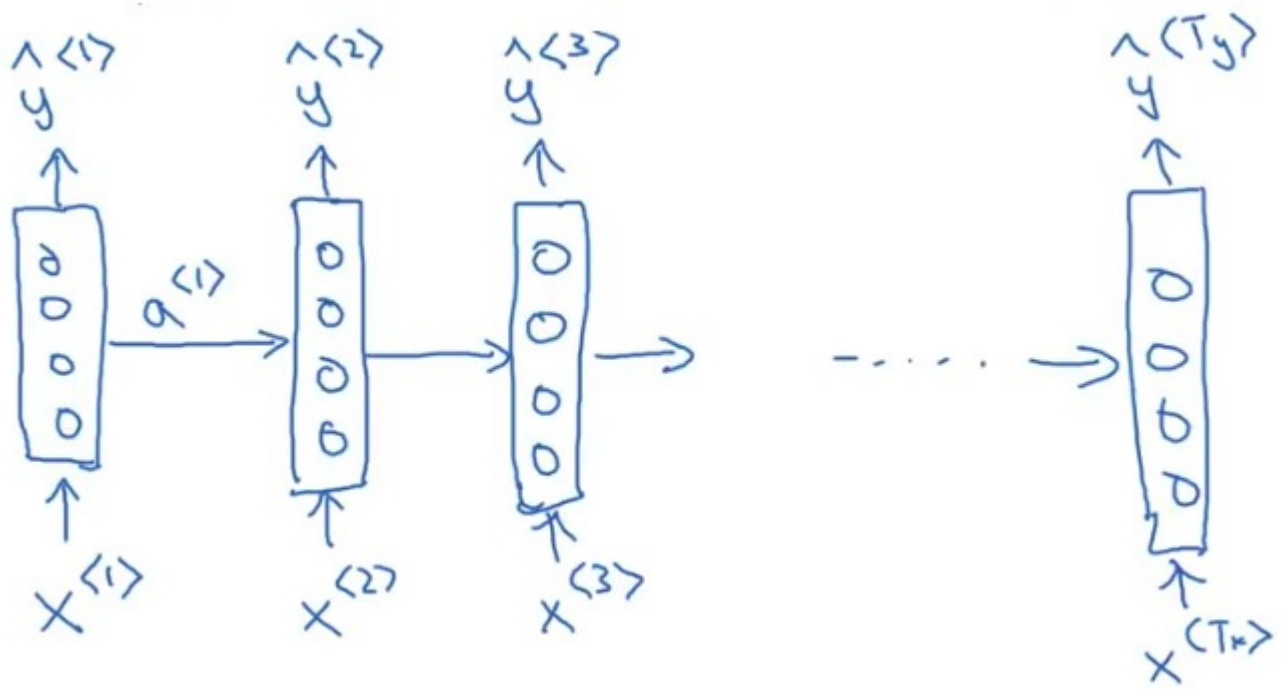
Notice: the same function and the same set of parameters are used at every time step.



Slide credit: F. Li & J. Johnson & S. Yeung

Rupam

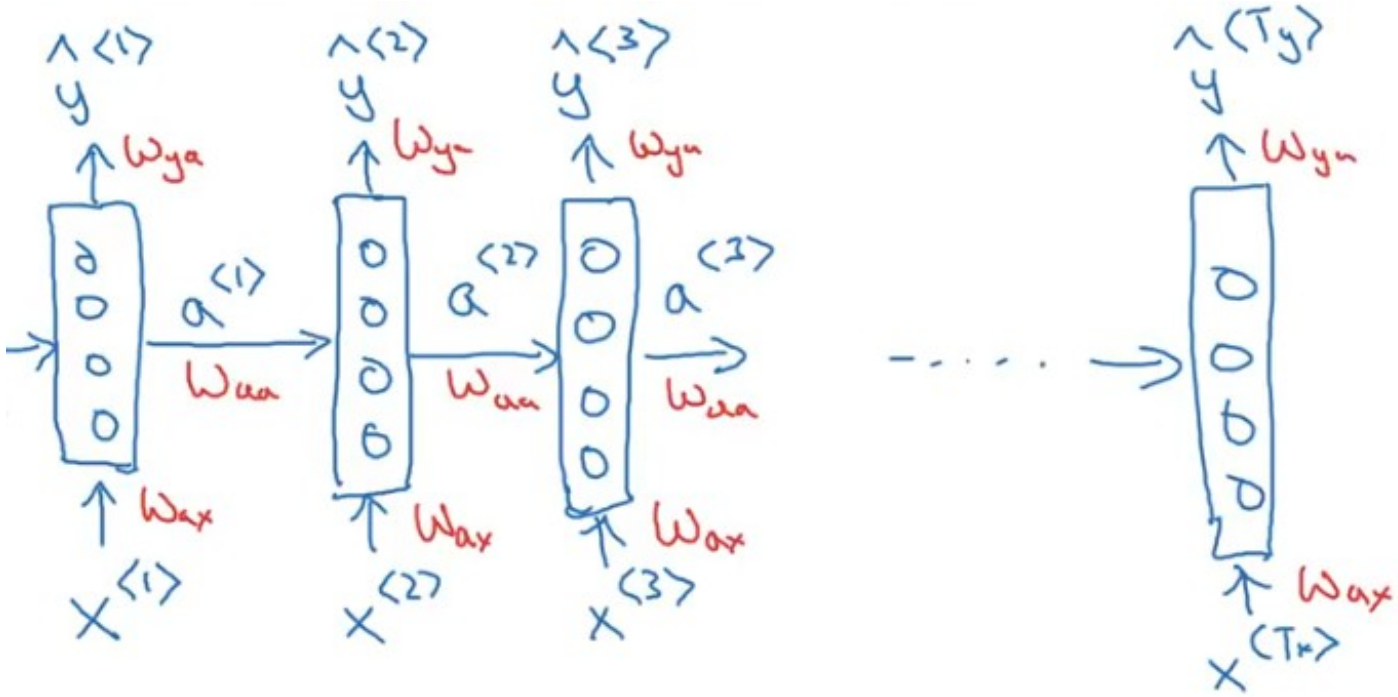
Example



Slide credit: Prof. A. Ng

Rupam

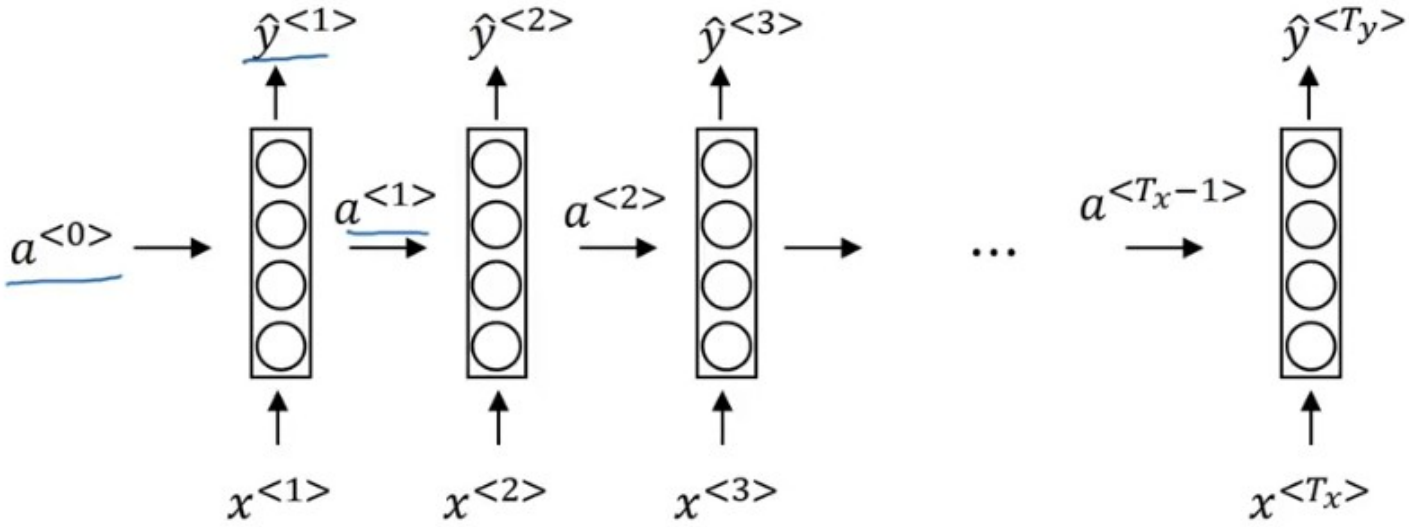
Example



Slide credit: Prof. A. Ng

Rupam

Example (forward propagation)



$a^{<0>} = \vec{0}$.

$a^{<1>} = g(W_{aa} a^{<0>} + W_{ax} x^{<1>} + b_a)$

$\hat{y}^{<1>} = g(W_{ya} a^{<1>} + b_y)$

Slide credit: Prof. A. Ng

References

- All the contents present in the slides are taken from various online resources. Due credit is given in the respective slides.These slides are used for *academic* purposes only.