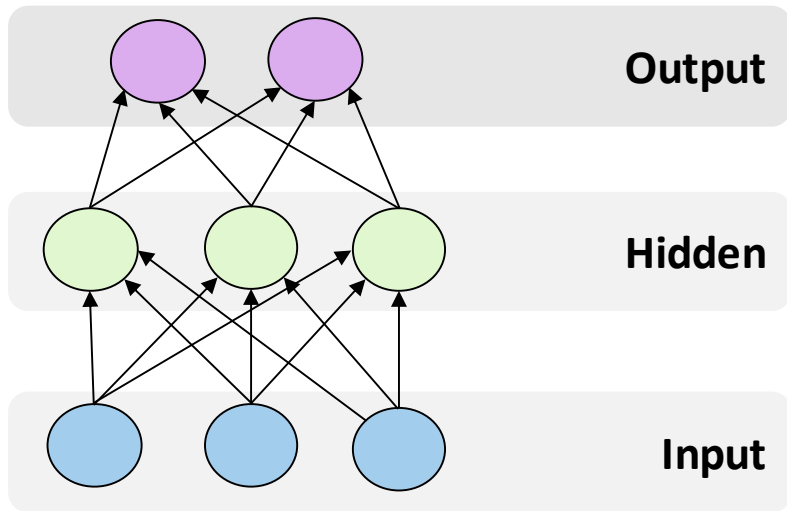


Recurrent Neural Networks

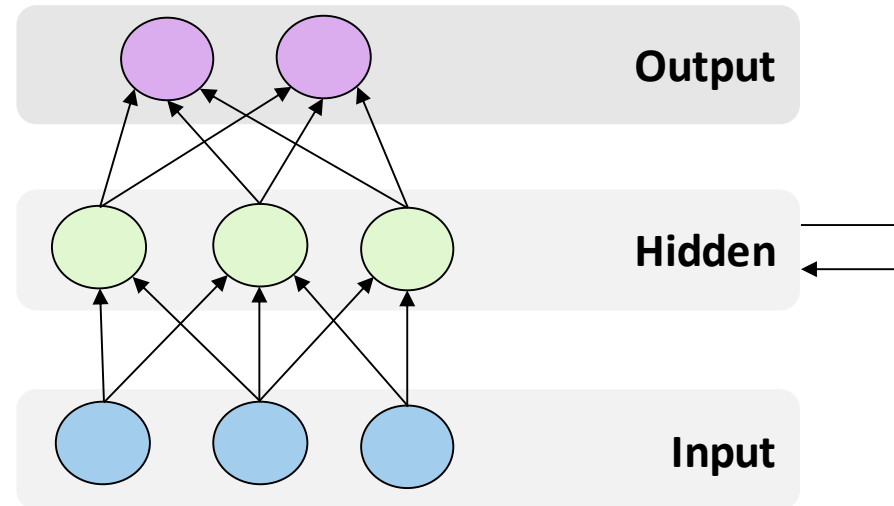
Aaron and Connor

Neural Network Architectures

- ❑ There are many variants on the Neural Networks we've studied in class.
- ❑ Any variant of a neural network idea is called a neural network architecture.
- ❑ These architectures can get pretty crazy!



Basic Neural Network



Recurrent Neural Network

When is it appropriate to use Recurrent Neural Networks?

□ Dealing with time-dependent inputs:

❖ Language processing

❖ Video processing

❖ Resource allocation

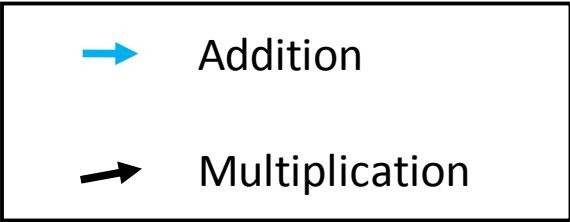
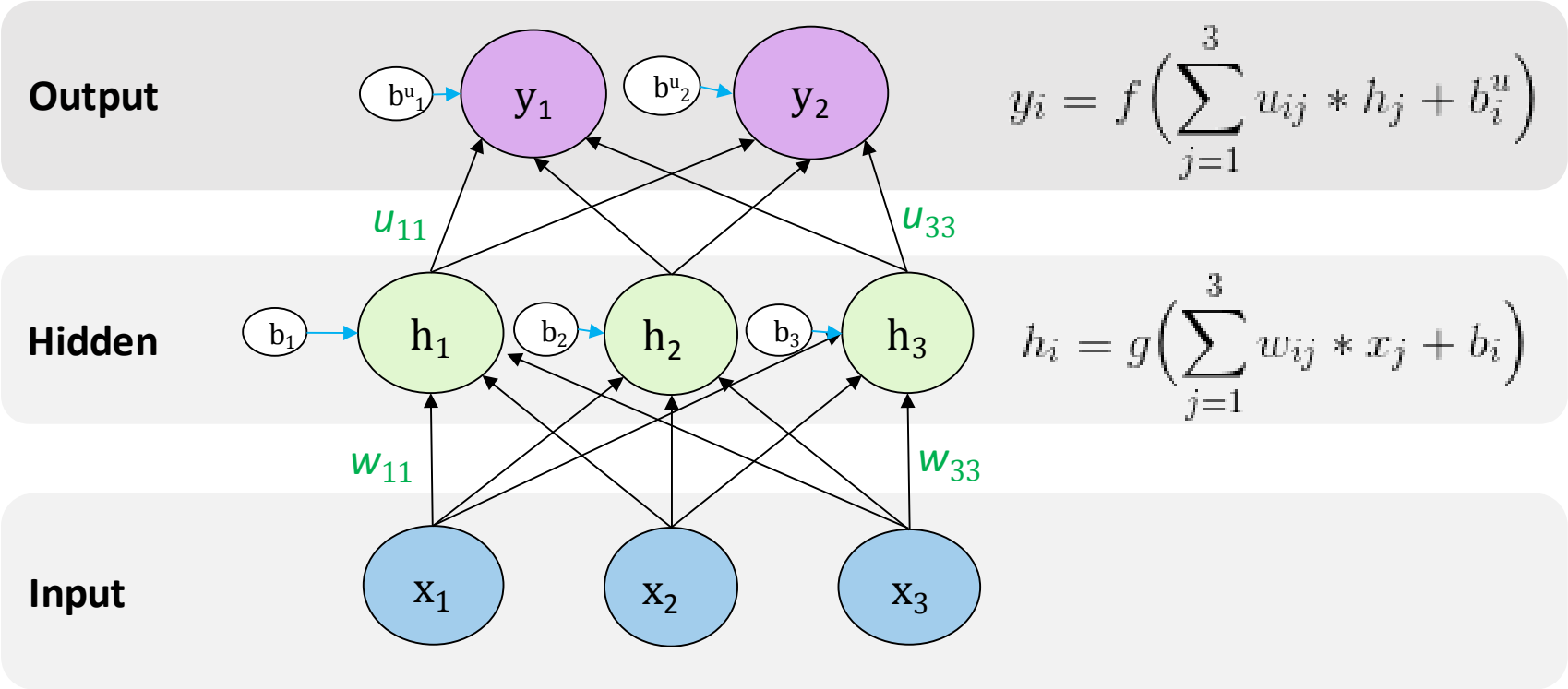
□ Any other sort of sequential features:

❖ Word prediction

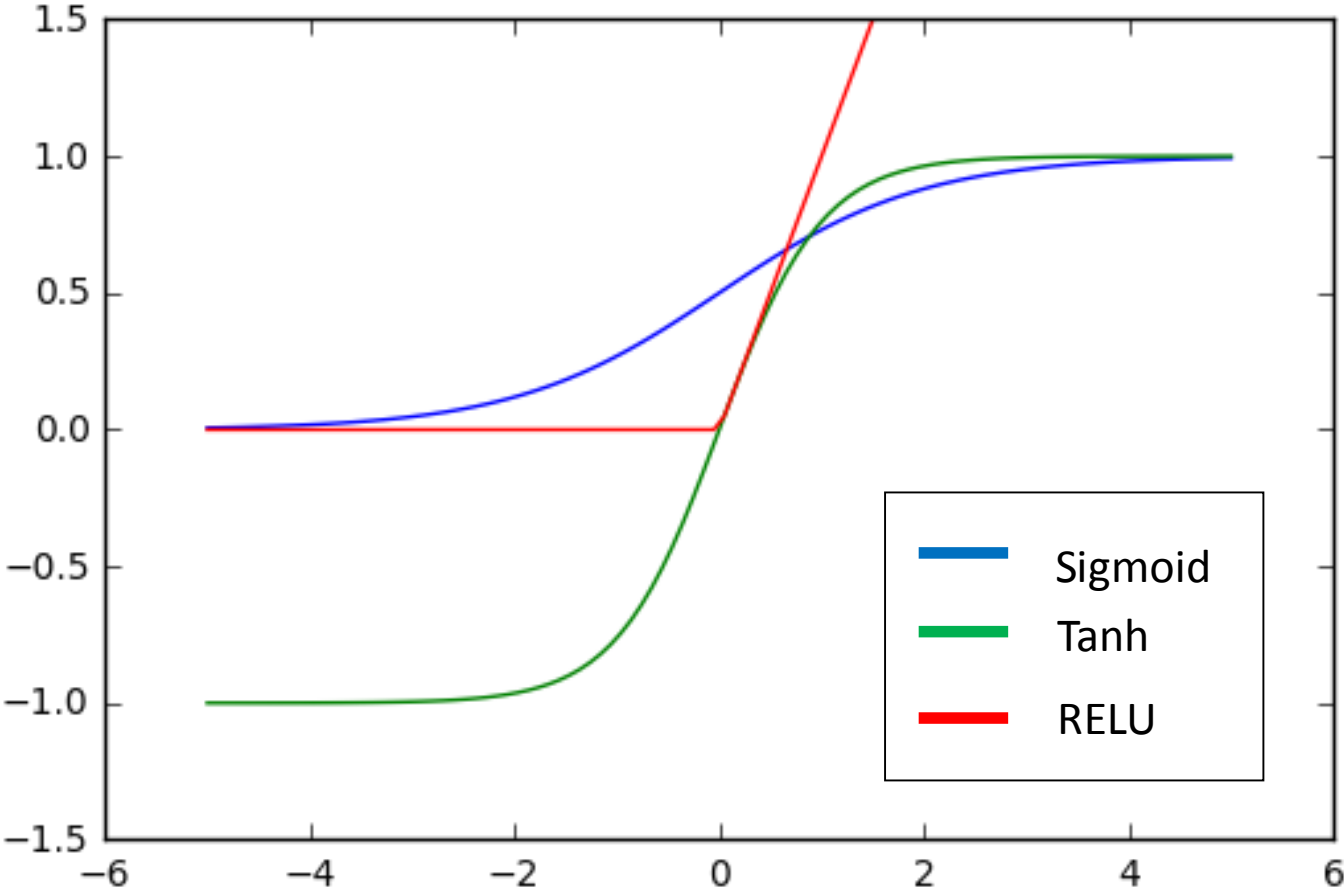
❖ Image processing

❖ Combinatorial optimization

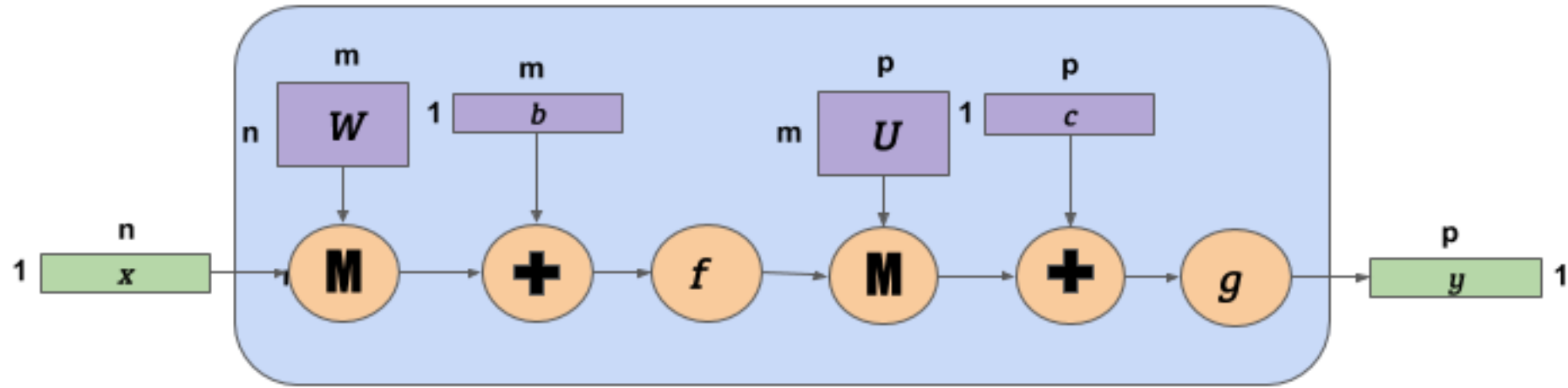
Review of Basic Single Layer Neural Network



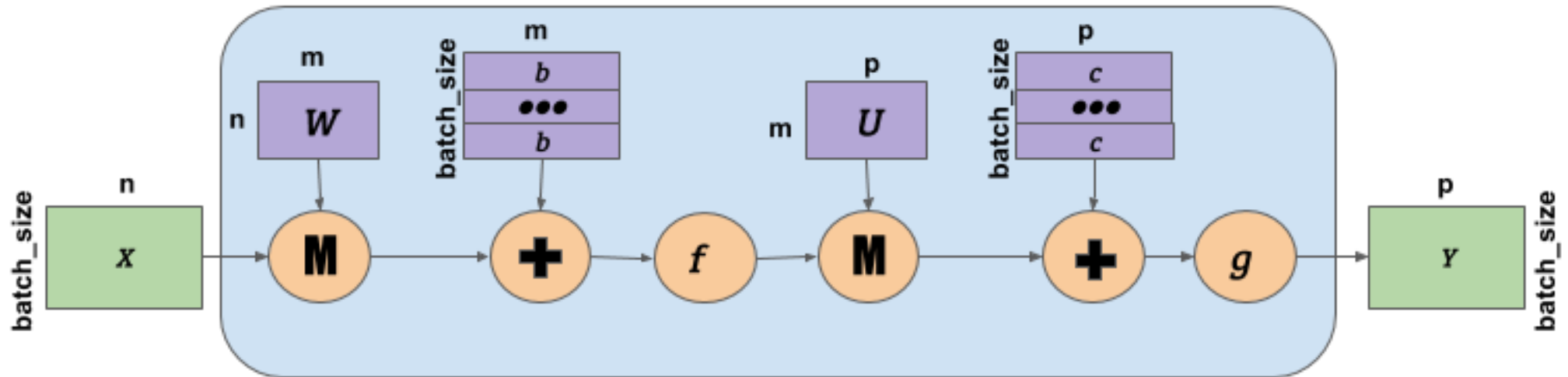
Common Activation Functions For Neural Nets



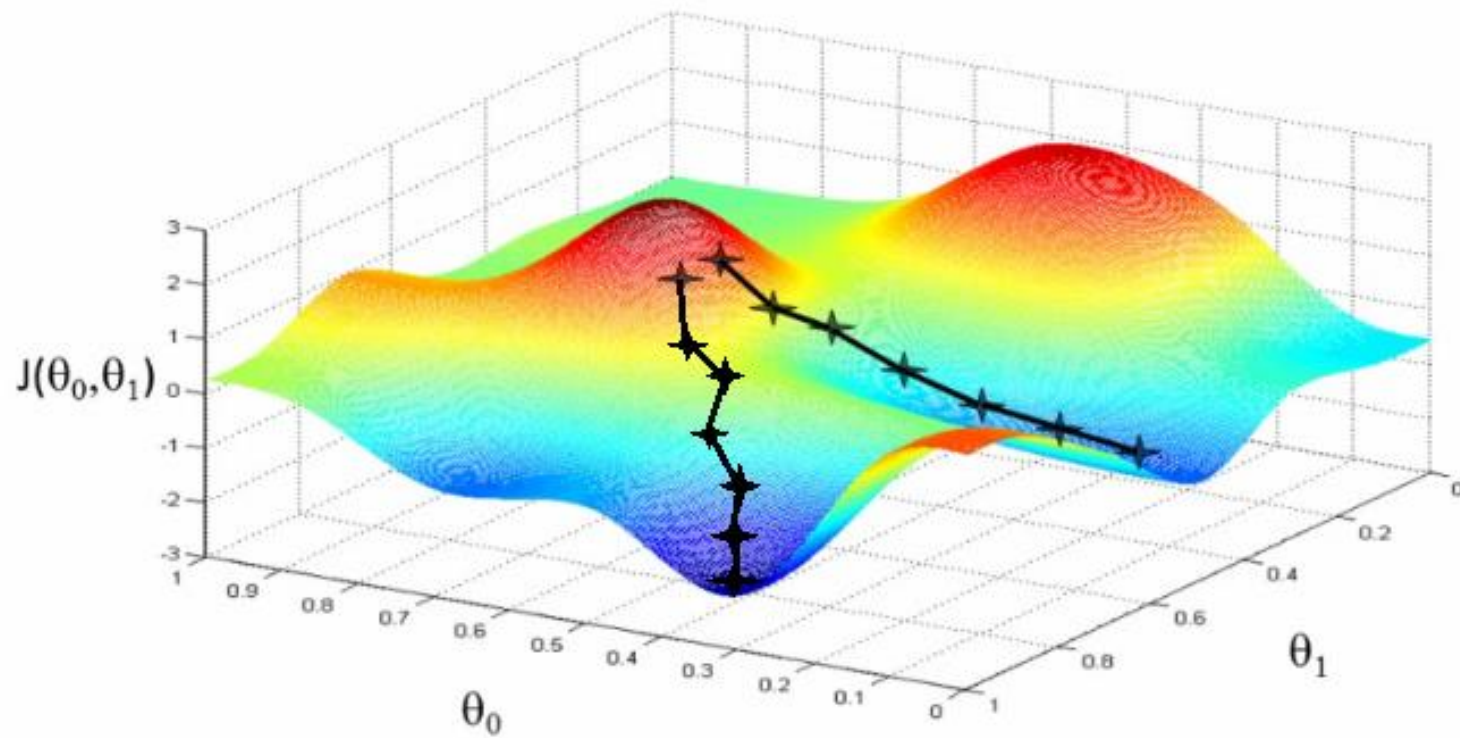
Computational Graph Representation of a Neural Network



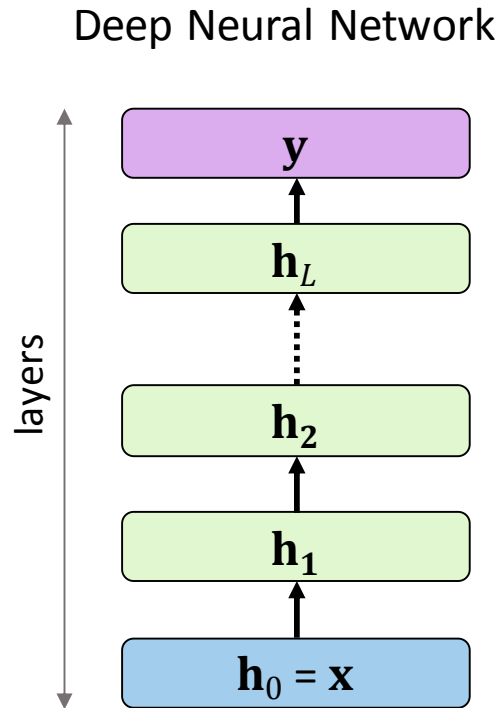
Neural Network Computational Graph



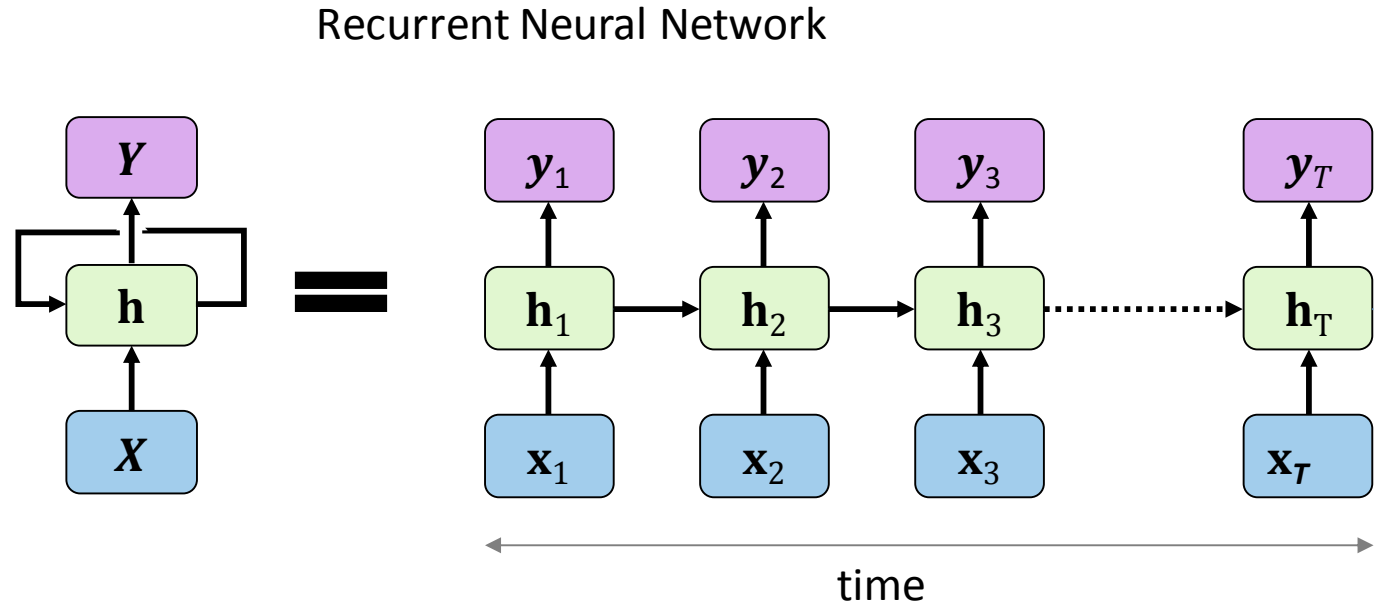
Training Neural Nets: Gradient Descent



Layer Representation of Neural Networks: Matrix Operations



$$\mathbf{y} = f(\mathbf{U}\mathbf{h}_L + \mathbf{b}^u)$$
$$\mathbf{h}_l = g(\mathbf{W}\mathbf{h}_{l-1} + \mathbf{b}^l)$$



$$\mathbf{y}_t = f(\mathbf{U}\mathbf{h}_T + \mathbf{b}^u)$$

$$\mathbf{h}_t = g(\mathbf{W}^h \mathbf{h}_{t-1} + \mathbf{W}^x \mathbf{x}_t + \mathbf{b}^h)$$

Examples

- ❑ Voice generation

- ❑ <https://www.youtube.com/watch?v=FsVSZpoUdSU>

- ❑ Image recognition

- ❑ <http://cs.stanford.edu/people/karpathy/deepimagesent/>

- ❑ Music composition

- ❑ <http://www.hexahedria.com/2015/08/03/composing-music-with-recurrent-neural-networks/>

- ❑ Language translation

- ❑ <http://104.131.78.120/>

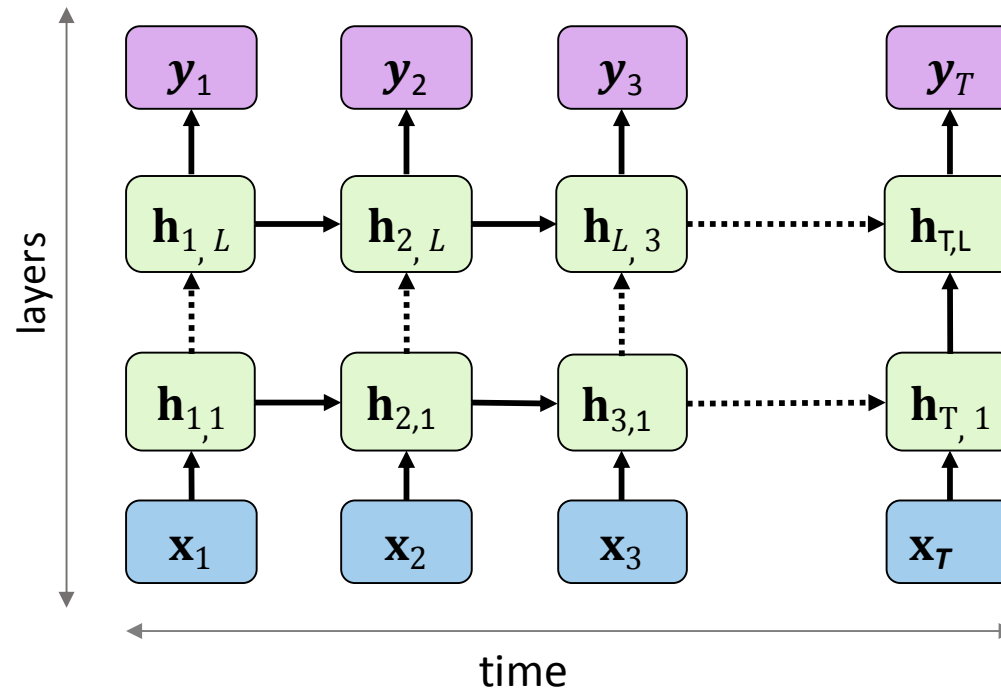
- ❑ Sequence prediction

- ❑ <https://www.youtube.com/watch?v=AJUgmj5NEwg>

- ❑ Handwriting generation

- ❑ <http://www.cs.toronto.edu/~graves/handwriting.html>

Deep Recurrent Neural Network

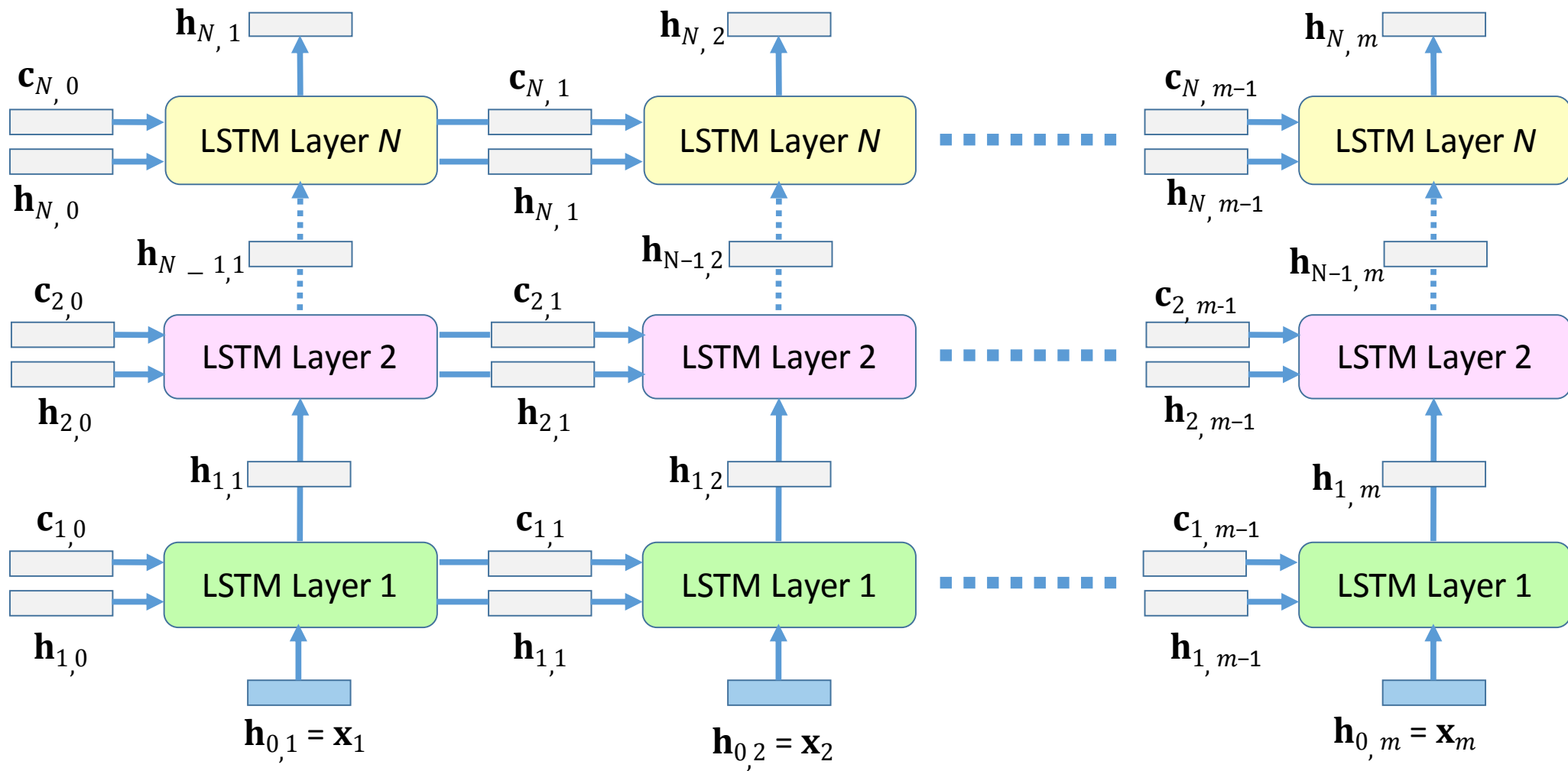


$$y_t = f(\mathbf{U}h_{T,L} + \mathbf{b}^u)$$

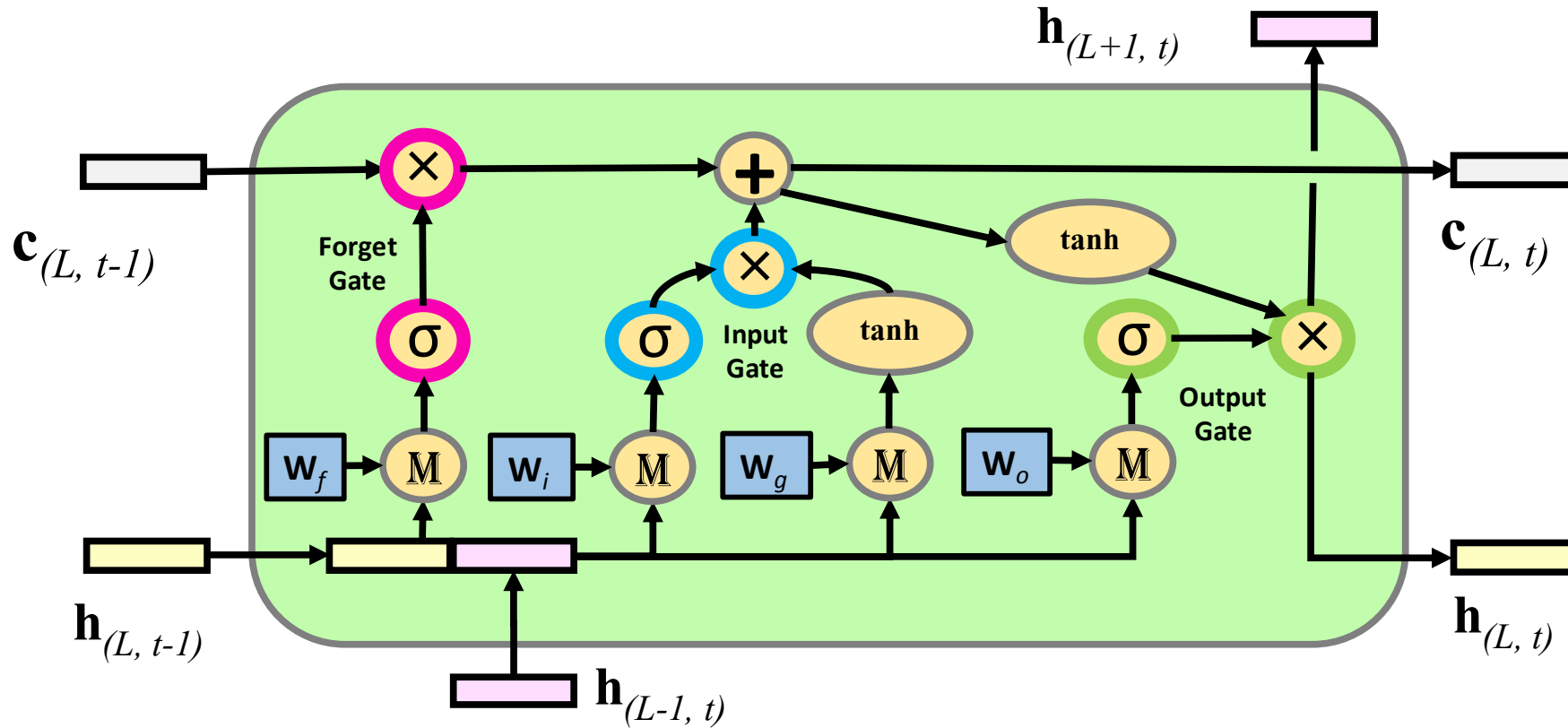
$$h_{t,l} = g(\mathbf{W}_l^h h_{t-1,l} + \mathbf{W}_l^x h_{t,l-1} + \mathbf{b}_l^h)$$

$$h_{t,0} = x_t$$

An Unrolled Deep Recurrent Neural Network with N LSTM Layers



LSTM Layer Computational Graph



- \times Elementwise multiplication
- $+$ Matrix Addition
- M Matrix multiplication
- σ Elementwise sigmoid
- \tanh Elementwise \tanh