DNN Discussion

講者：Isaac

▸ Why DNN Fail in early day

Why DNN Fail in early day

- Some issue that make DNN fail in early day
  - Hardware issue
    - Doesn't have fast enough GPU like today
  - Data issue
    - Hard to collect large amount of data
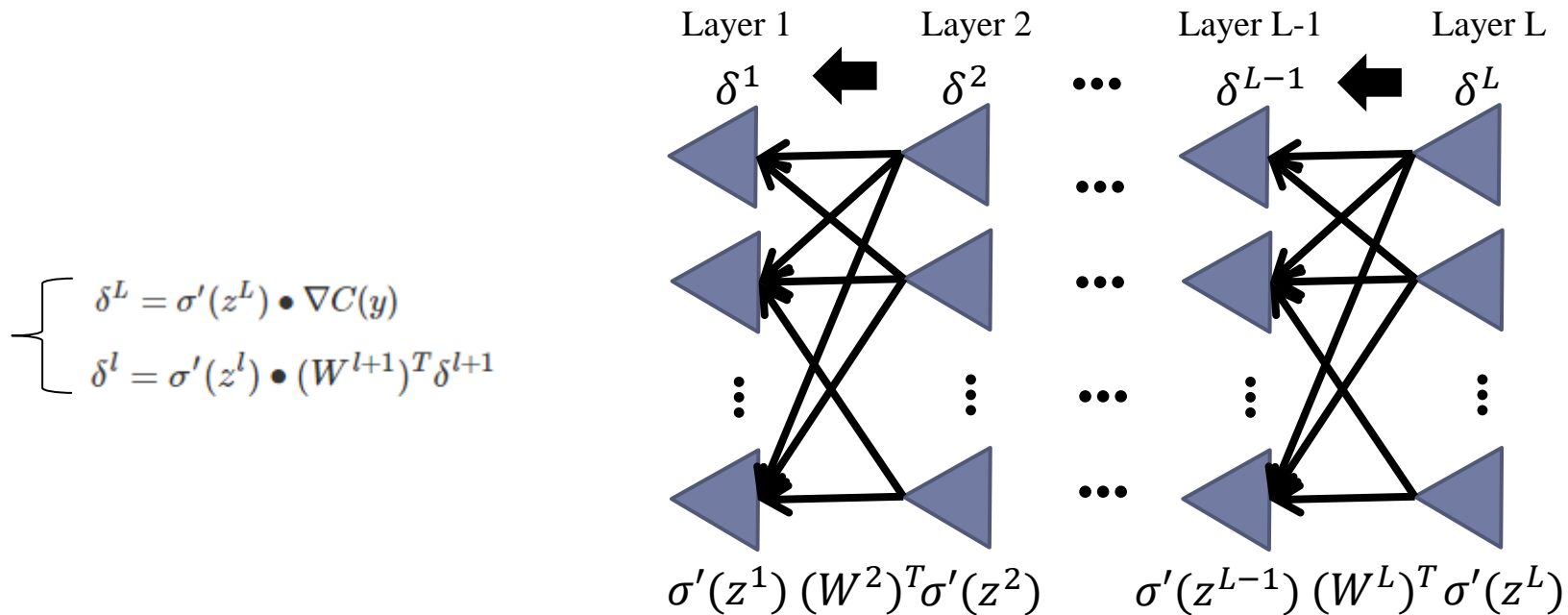 - Activation Function
    - Use sigmoid as activation function

Activation Function

▸ Give NN nonlinearity property

▸ Can be regarded as "ON" (1) or "OFF" (0)

Input layer   Layer 1   Layer 2   Layer L   Output layer
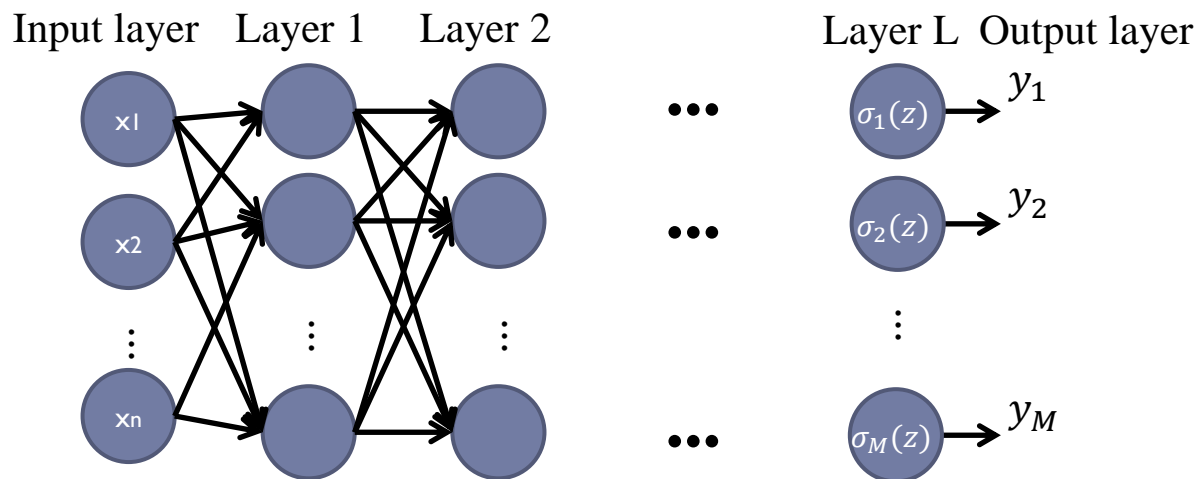
$$y = f(x) = \sigma(W^L \dots \sigma(W^2 \sigma(W^1 X + b^1) + b^2) \dots + b^L)$$

Backpropagation

$$\delta^L = \sigma'(z^L) \bullet \nabla C(y)$$
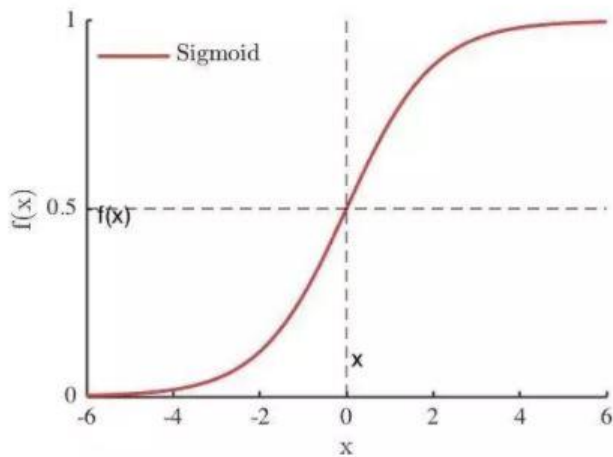
$$\delta^l = \sigma'(z^l) \bullet (W^{l+1})^T \delta^{l+1}$$

Backpropagation = many multiplication of derivative activation function

Sigmoid Function

▸ Maximinum of slope is ¼
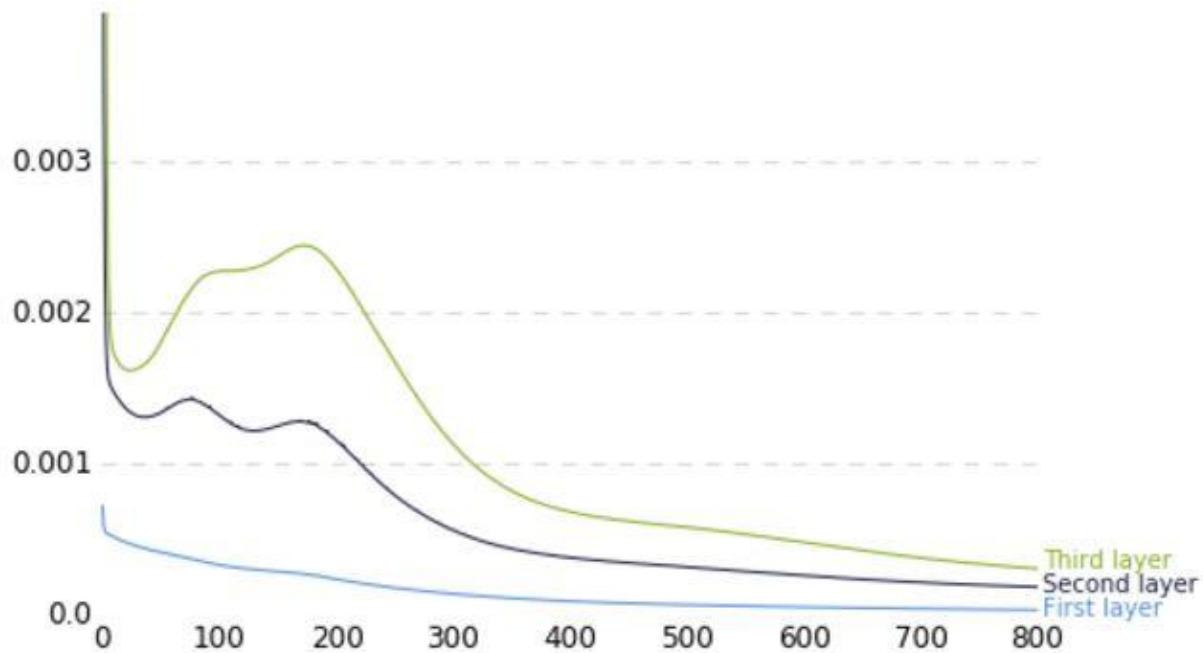
$$f(x) = \frac{1}{1 + e^{-x}}$$

RELU

▸ slope is

  ▸ 1 when x >= 0

  ▸ 0 when x < 0
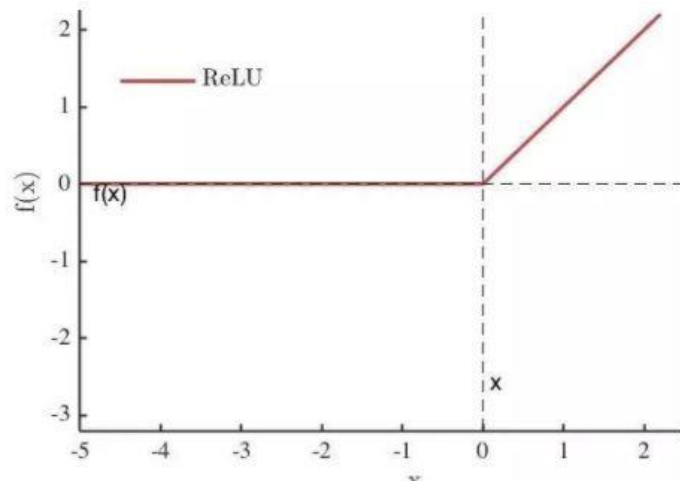
$$y = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Gradient vanishing illustration