22101622

1

Suppose they are n layered.

For each layer, identity activation function $f(x) = x$; input = x.

for a neuron, $y = f(w_i x_i + b)$

Let say, for layer 1,

$y_1 = f(w_1 x_0 + b_1) = w_1 x_0 + b_1$ [$f(x) = x$]

~~1-2 : $y_2 = f(w_2 x_2 + b_2) = w_1 (w_2 x_0 + b_2) + b_1$~~

~~$= w_1 w_2 x + w_1 b_2 + b_1$~~

1-2 : $y_2 = f(w_2 y_1 + b_2) = w_2 y_1 + b_2$

$= w_2 (w_1 x + b_1) + b_2$

$= \underbrace{w_1 w_2}_{w} x + \underbrace{w_2 b_1 + b_2}_{b}$

1-3 : $y_3 = f(w_3 y_2 + b_3) = w_3 y_2 + b_3$

$= \underbrace{w_3 w_1 w_2}_{w} x + \underbrace{w_3 w_2 b_1 + w_3 b_2 + b_3}_{b}$

. . . . .

n: $y_n = f(w_n y_{n-1} + b_n) =$

$$\underbrace{w_n w_{n-1} \cdots w_1 x}_{W} + \underbrace{w_n w_{n-1} \cdots w_2 b_1 + w_n b_{n-1} + \cdots +}_{b \text{ (bias)}}$$

$y_n = Wx + b$

So, $f(x) = x$ identity & reduces a ~~multilayered~~ pawered multilayer network

~~network~~ it to a single layer network (proved)