

Forward Propagation Algorithm

Step-by-Step Process

1 Input Layer
 $h^{(0)} = x$



2 Linear Transformation
 $z^{(1)} = W^{(1)}h^{(1-1)} + b^{(1)}$



3 Activation Function
 $h^{(1)} = \sigma(z^{(1)})$



4 Repeat for All Layers
 $l = 1, 2, \dots, L$



5 Output Layer
 $\hat{y} = h^{(L)}$

Pseudocode

Forward Propagation

```
# Input: x, weights W, biases b
function forward_propagation(x):
    h = x # Initialize with input

    for l = 1 to L:
        # Linear combination
        z = W[l] * h + b[l]

        # Apply activation
        h = activation(z)

    return h # Final output  $\hat{y}$ 
```

💡 Example: 3-Layer Network

$x \rightarrow [W^1, b^1] \rightarrow \sigma \rightarrow h^1 \rightarrow [W^2, b^2] \rightarrow \sigma \rightarrow h^2 \rightarrow [W^3, b^3] \rightarrow \sigma \rightarrow \hat{y}$



Sequential Flow

Data flows layer by layer from input to output



Two Operations

Each layer: (1) linear transform (2) activation



Matrix Operations

Efficient computation using matrix multiplication



Numerical Example: Step-by-Step Calculation

2-Layer Network with Sigmoid Activation ($\sigma(z) = 1/(1+e^{-z})$)



Initial Setup

Input: $x = [1.0, 2.0]^T$

Layer 1 Weights: $W^{(1)} = \begin{bmatrix} 0.5 & 0.3 \\ 0.2 & 0.4 \end{bmatrix}$

Layer 1 Bias: $b^{(1)} = [0.1, 0.2]^T$

Layer 2 Weights: $W^{(2)} = [0.6, 0.7]$

Layer 2 Bias: $b^{(2)} = [0.3]$



Layer 1: Hidden Layer

Step 1: Linear Transformation

$$z^{(1)} = W^{(1)}x + b^{(1)}$$

$$z_1^{(1)} = 0.5(1.0) + 0.3(2.0) + 0.1 = 0.5 + 0.6 + 0.1 = 1.2$$

$$z_2^{(1)} = 0.2(1.0) + 0.4(2.0) + 0.2 = 0.2 + 0.8 + 0.2 = 1.2$$

$$z^{(1)} = [1.2, 1.2]^T$$



Step 2: Activation Function

$$h^{(1)} = \sigma(z^{(1)})$$

$$h_1^{(1)} = \sigma(1.2) = 1/(1+e^{-1.2}) \approx 0.769$$

$$h_2^{(1)} = \sigma(1.2) = 1/(1+e^{-1.2}) \approx 0.769$$

$$\mathbf{h}^{(1)} = [0.769, 0.769]^T$$

Layer 2: Output Layer

Step 3: Linear Transformation

$$\mathbf{z}^{(2)} = \mathbf{w}^{(2)}\mathbf{h}^{(1)} + \mathbf{b}^{(2)}$$

$$z^{(2)} = 0.6(0.769) + 0.7(0.769) + 0.3$$

$$z^{(2)} = 0.461 + 0.538 + 0.3 = \mathbf{1.299}$$

$$\mathbf{z}^{(2)} = [1.299]$$



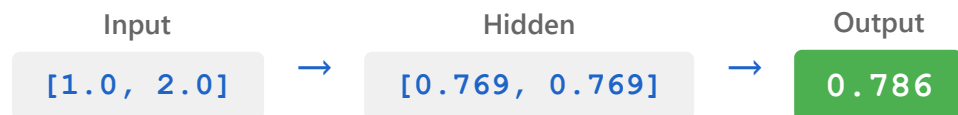
Step 4: Final Activation

$$\hat{y} = \mathbf{h}^{(2)} = \sigma(\mathbf{z}^{(2)})$$

$$\hat{y} = \sigma(1.299) = 1/(1+e^{-1.299}) \approx \mathbf{0.786}$$

$$\hat{y} = \mathbf{0.786}$$

Summary



💡 This output represents the network's prediction for the given input after forward propagation through all layers.