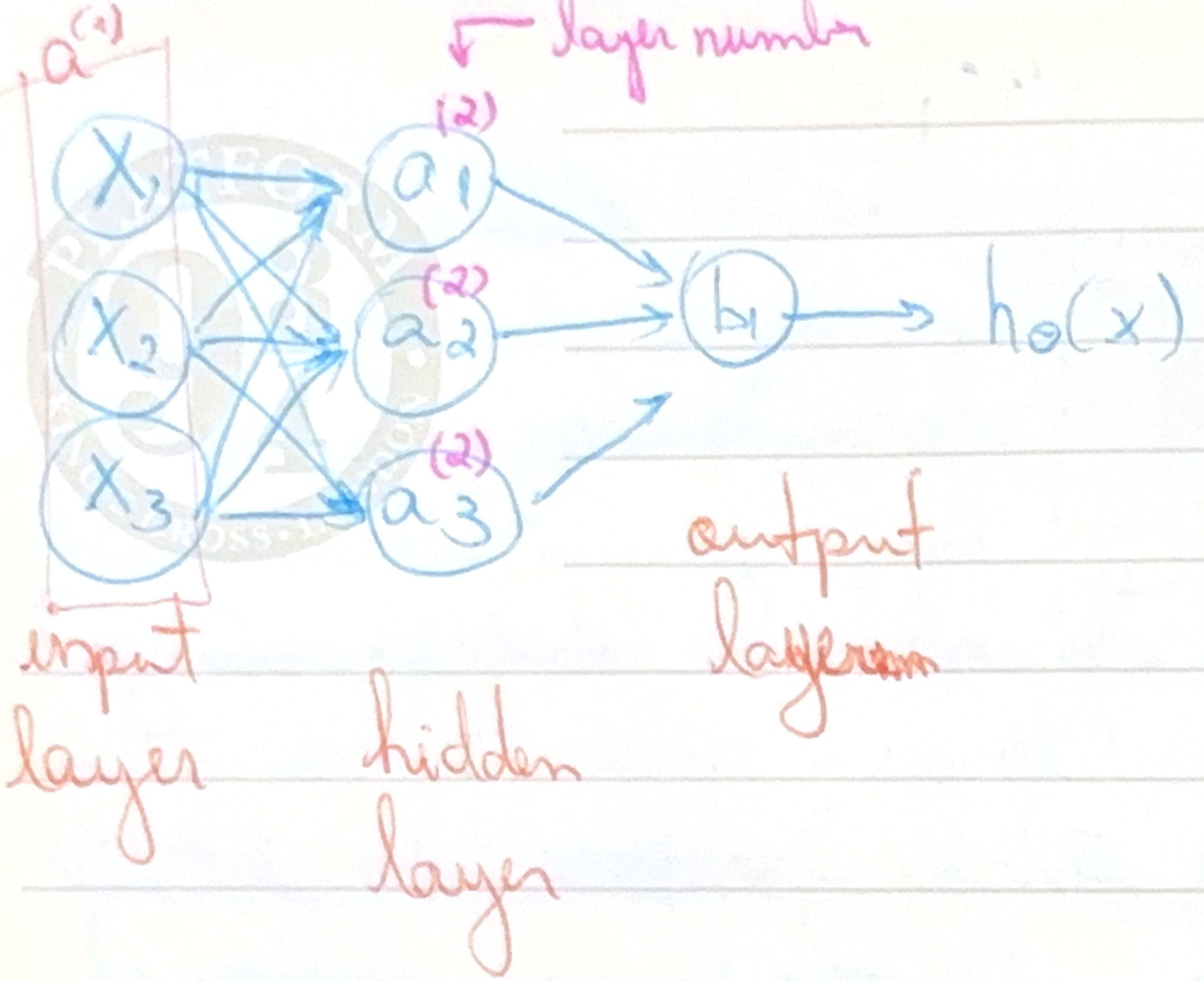

$\Theta^{(j)}$ = matrix of weights controlling function mapping from layer j to layer j+1

Exemplo:

$$\Theta^{(1)} = \begin{bmatrix} \Theta_{10}^{(1)} & \Theta_{11}^{(1)} & \Theta_{12}^{(1)} & \Theta_{13}^{(1)} \\ \Theta_{20}^{(1)} & \Theta_{21}^{(1)} & \Theta_{22}^{(1)} & \Theta_{23}^{(1)} \\ \Theta_{30}^{(1)} & \Theta_{31}^{(1)} & \Theta_{32}^{(1)} & \Theta_{33}^{(1)} \end{bmatrix}$$

dim 3x4 remetente
destinatário

$\Theta_{jK}^{(i)}$ → layer onde o parâmetro é aplicado

"DESTINATÁRIO"
j: posição do neurônio de destino dentro da camada seguinte (i+1)

"REMETENTE"
K: posição do neurônio de partida dentro da camada (i)

second layer

$a_1^{(2)} = g(z_1^{(2)})$

1: first neuron

$z_1^{(2)}$: $\Theta^T x$ for the first neuron of the second layer. This $\Theta^T x$ will be the argument for the sigmoid function $g(z) = \frac{1}{1+e^{-z}}$

IMPLEMENTAÇÃO COM VETORES:

$z^{(2)} = \Theta^{(1)} \cancel{x} \to a^{(1)}$

$a^{(2)} = g(z^{(2)})$

$\mathbb{R}^3$; $\mathbb{R}^3$

Add $a_0^{(2)} = 1 \to a_2 \in \mathbb{R}^4$

$z^{(3)} = \Theta^{(2)} a^{(2)}$

$h_\theta(x) = a^{(3)} = g(z^{(3)})$