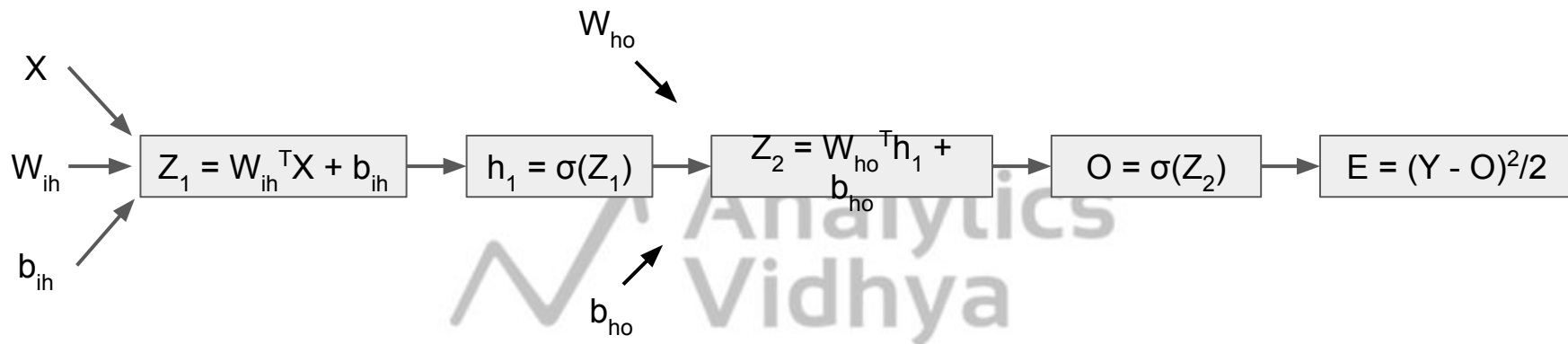


Backward Propagation: Matrix Form

Understanding Backward Propagation Mathematically



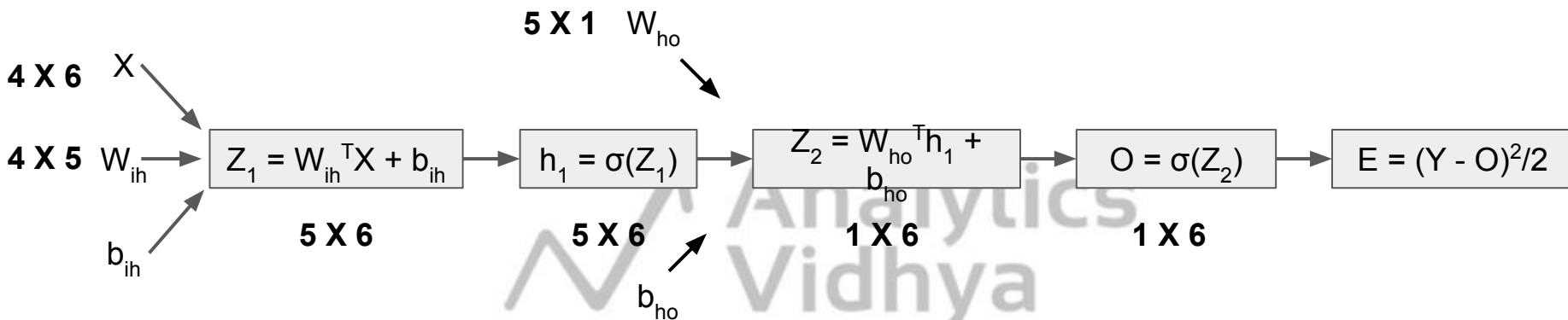
$$\frac{dE}{db_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{db_{ho}}$$

$$\frac{dE}{dW_{ih}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dh_1} * \frac{dh_1}{dZ_1} * \frac{dZ_1}{dW_{ih}}$$

$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

$$\frac{dE}{db_{ih}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dh_1} * \frac{dh_1}{dZ_1} * \frac{dZ_1}{db_{ih}}$$

Understanding Backward Propagation Mathematically



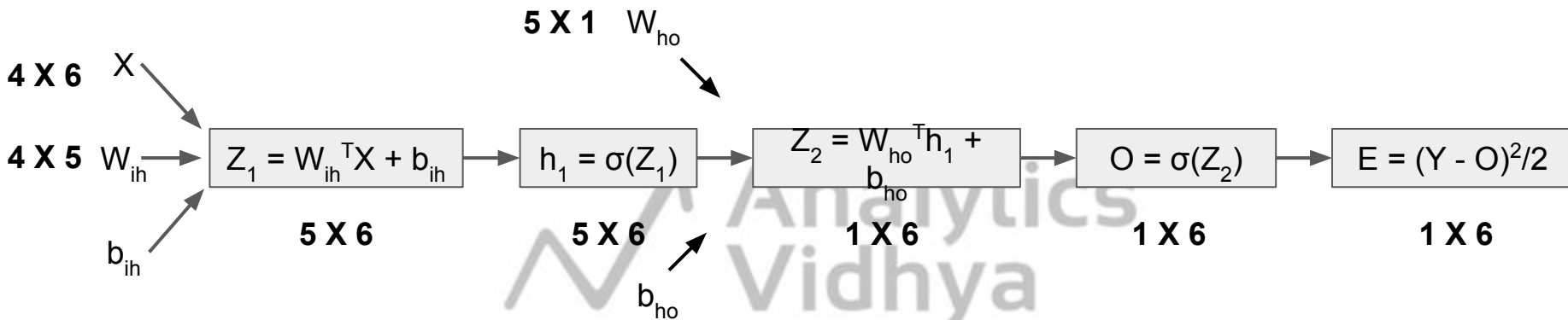
$$\frac{dE}{db_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{db_{ho}}$$

$$\frac{dE}{dW_{ih}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dh_1} * \frac{dh_1}{dZ_1} * \frac{dZ_1}{dW_{ih}}$$

$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

$$\frac{dE}{db_{ih}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dh_1} * \frac{dh_1}{dZ_1} * \frac{dZ_1}{db_{ih}}$$

Understanding Backward Propagation Mathematically



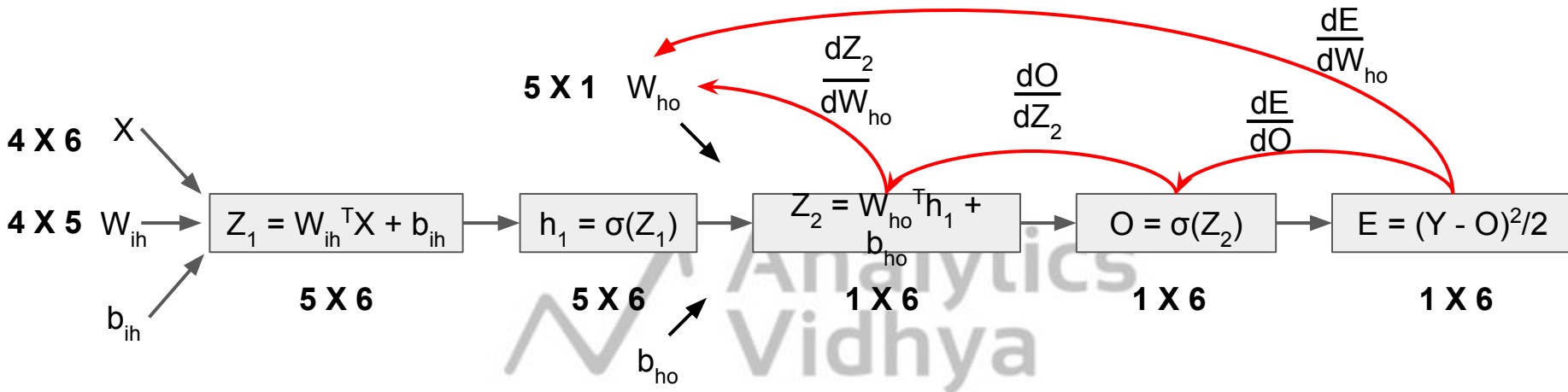
$$\frac{dE}{db_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{db_{ho}}$$

$$\frac{dE}{dW_{ih}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dh_1} * \frac{dh_1}{dZ_1} * \frac{dZ_1}{dW_{ih}}$$

$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

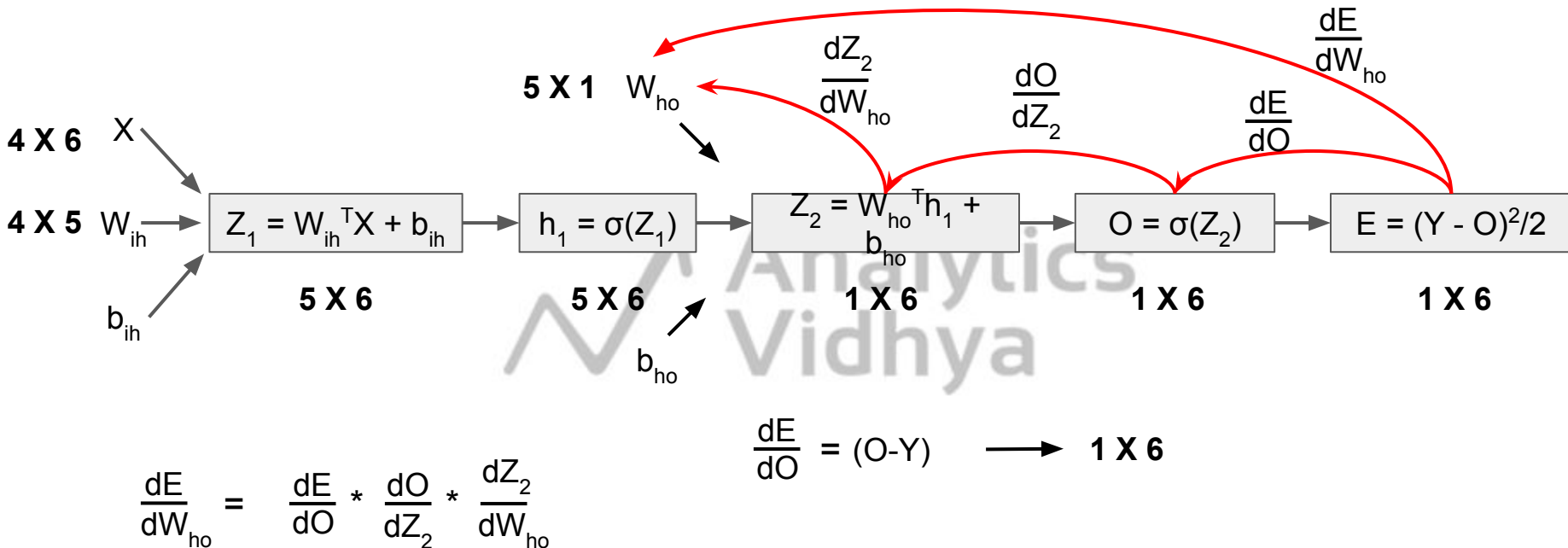
$$\frac{dE}{db_{ih}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dh_1} * \frac{dh_1}{dZ_1} * \frac{dZ_1}{db_{ih}}$$

Understanding Backward Propagation Mathematically

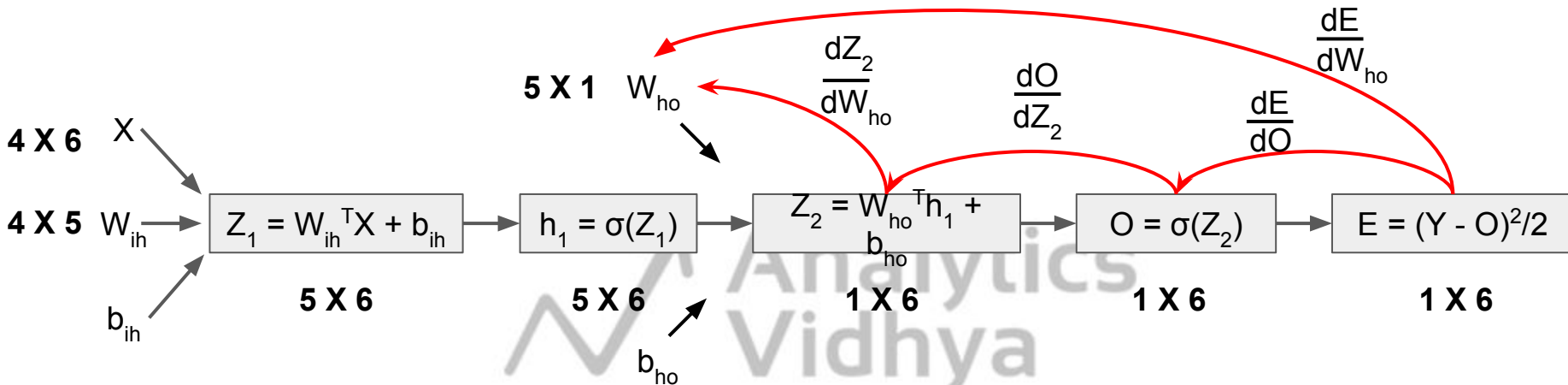


$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

Understanding Backward Propagation Mathematically



Understanding Backward Propagation Mathematically

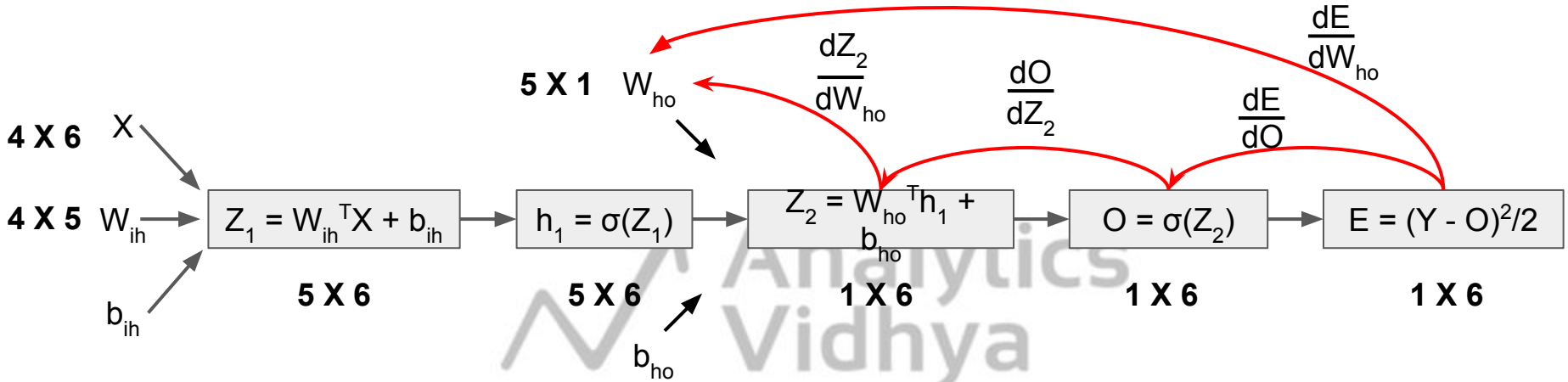


$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

$$\frac{dE}{dO} = (O - Y) \longrightarrow 1 \times 6$$

$$\frac{dO}{dZ_2} = O(1 - O) \longrightarrow 1 \times 6$$

Understanding Backward Propagation Mathematically



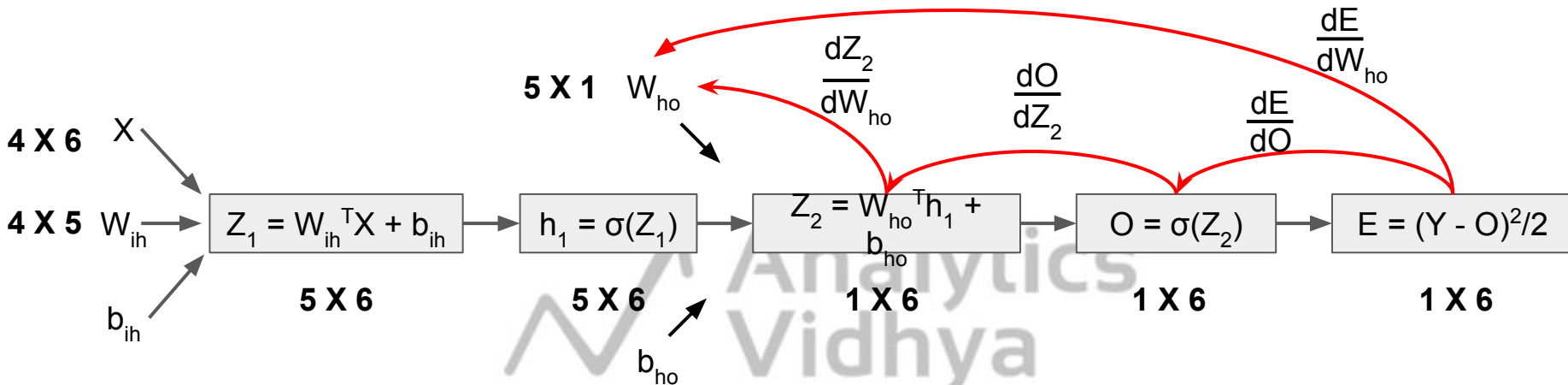
$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

$$\frac{dE}{dO} = (O - Y) \longrightarrow 1 \times 6$$

$$\frac{dO}{dZ_2} = O(1 - O) \longrightarrow 1 \times 6$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically



$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

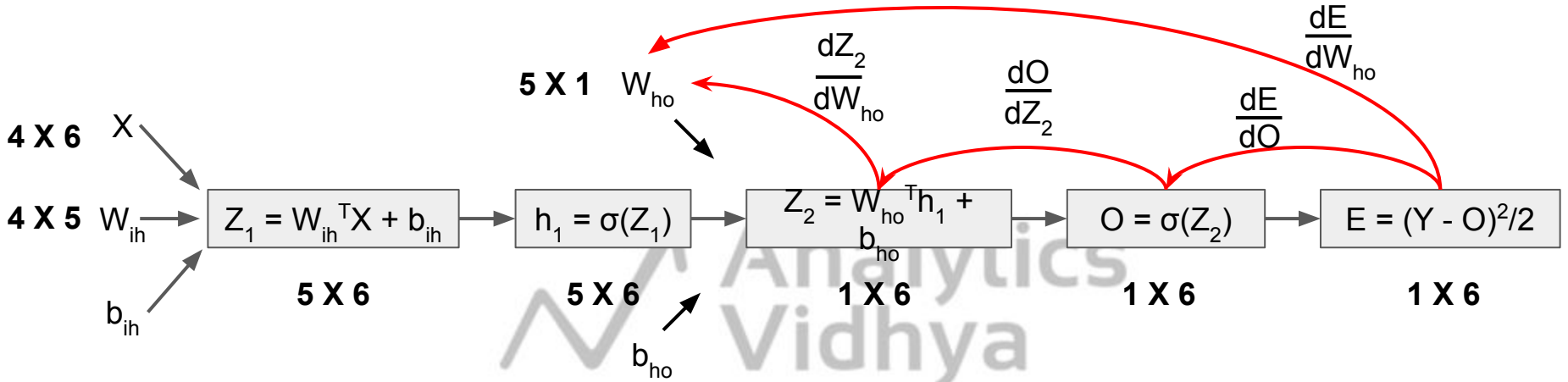
$$W_{ho} = W_{ho} - \alpha * \frac{dE}{dW_{ho}}$$

$$\frac{dE}{dO} = (O - Y) \longrightarrow 1 \times 6$$

$$\frac{dO}{dZ_2} = O(1 - O) \longrightarrow 1 \times 6$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically



$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

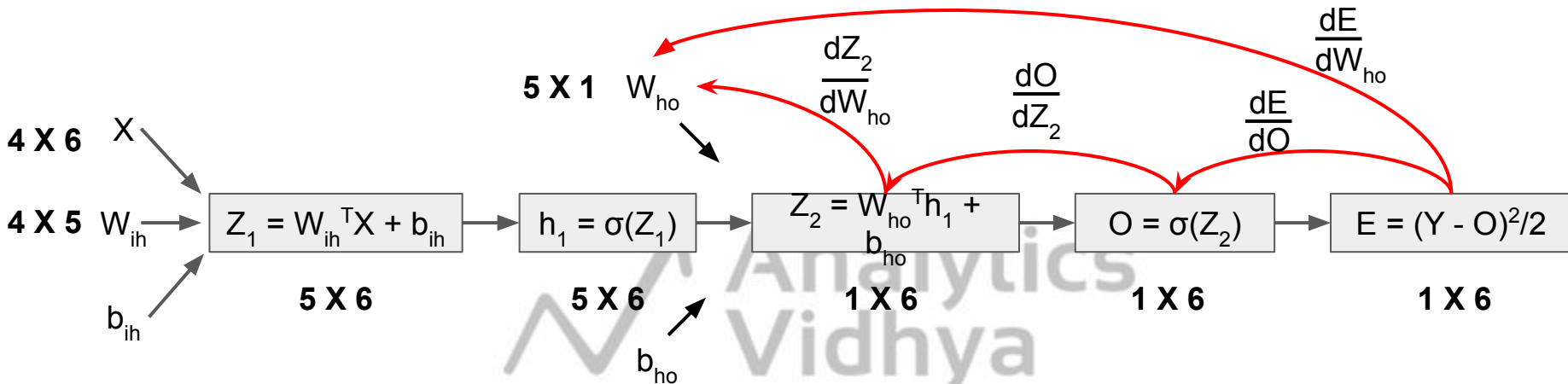
$$W_{ho} = W_{ho} - \alpha * \frac{dE}{dW_{ho}} \rightarrow 5 \times 1$$

$$\frac{dE}{dO} = (O - Y) \rightarrow 1 \times 6$$

$$\frac{dO}{dZ_2} = O(1 - O) \rightarrow 1 \times 6$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \rightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically



$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

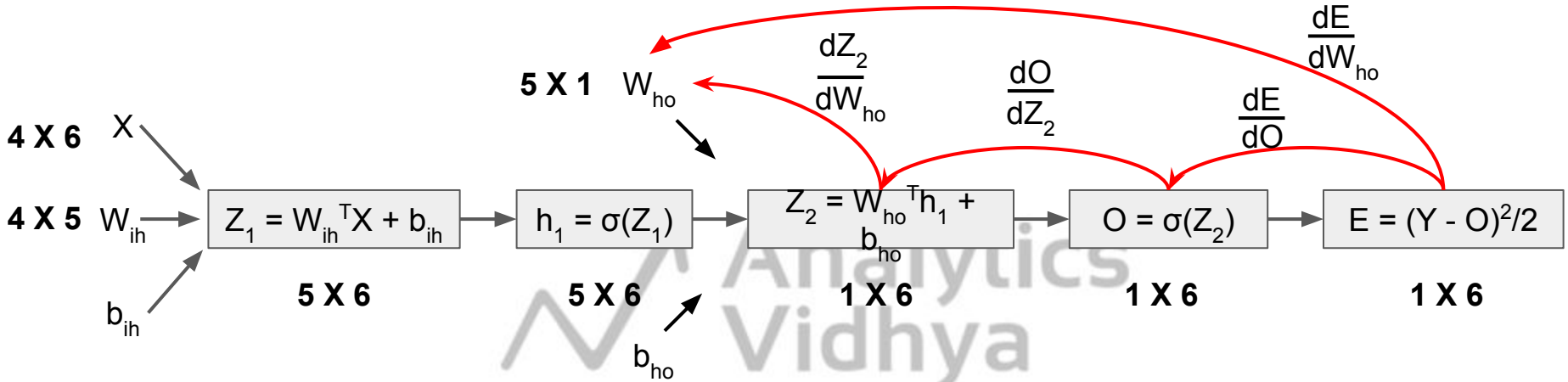
$$\frac{dE}{dW_{ho}} = (1 \times 6) * (1 \times 6) * (5 \times 6)$$

$$\frac{dE}{dO} = (O - Y) \longrightarrow 1 \times 6$$

$$\frac{dO}{dZ_2} = O(1 - O) \longrightarrow 1 \times 6$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically



$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

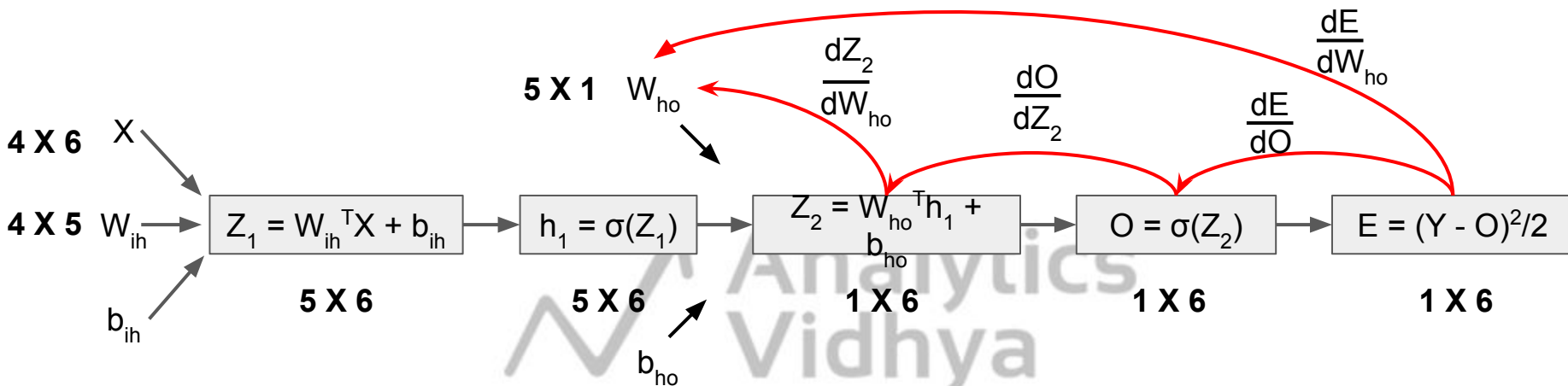
$$\frac{dE}{dW_{ho}} = \text{Can not perform matrix multiplication}$$

$$\frac{dE}{dO} = (O - Y) \longrightarrow 1 \times 6$$

$$\frac{dO}{dZ_2} = O(1 - O) \longrightarrow 1 \times 6$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically



$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

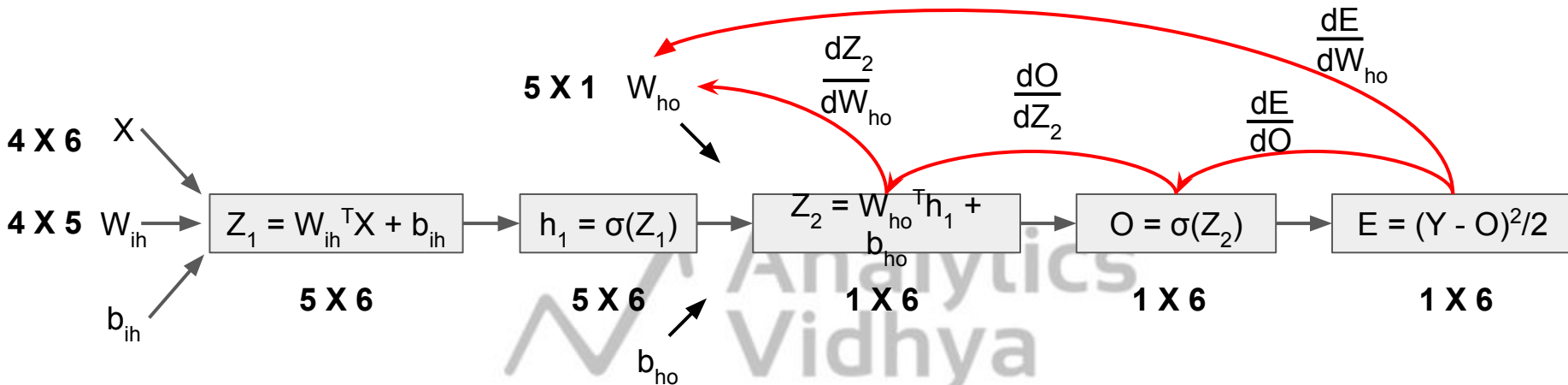
$$\frac{dE}{dW_{ho}} = (5 \times 6)$$

$$\frac{dE}{dO} = (O - Y) \longrightarrow 1 \times 6$$

$$\frac{dO}{dZ_2} = O(1 - O) \longrightarrow 1 \times 6$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically



$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

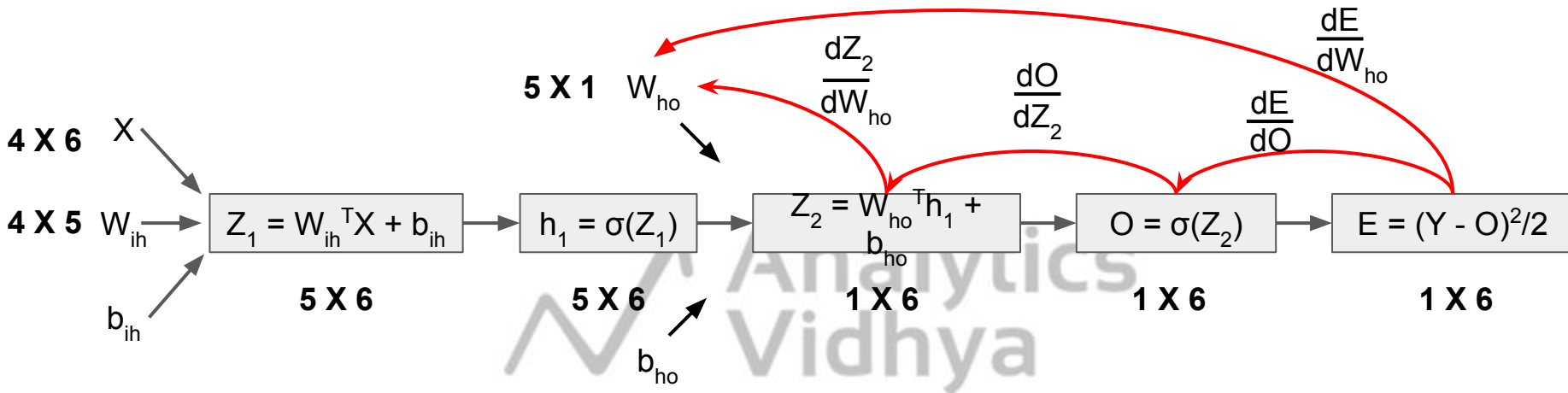
$$\frac{dE}{dO} * \frac{dO}{dZ_2} = (1 \times 6) * (1 \times 6) = (1 \times 6)$$

$$\frac{dE}{dO} = (O - Y) \longrightarrow 1 \times 6$$

$$\frac{dO}{dZ_2} = O(1 - O) \longrightarrow 1 \times 6$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically

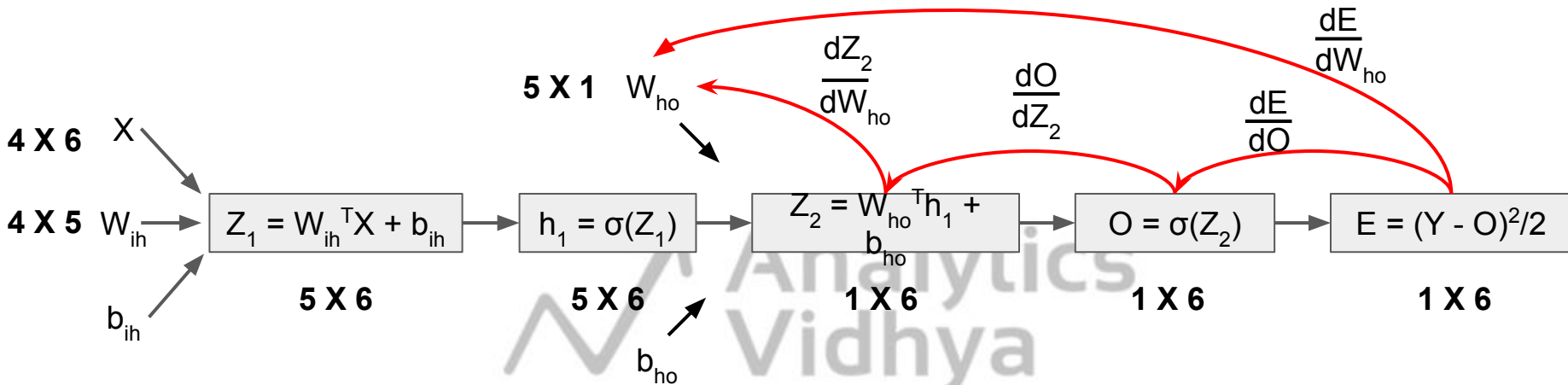


$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

$$\frac{dE}{dO} * \frac{dO}{dZ_2} = (1 \times 6)$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically

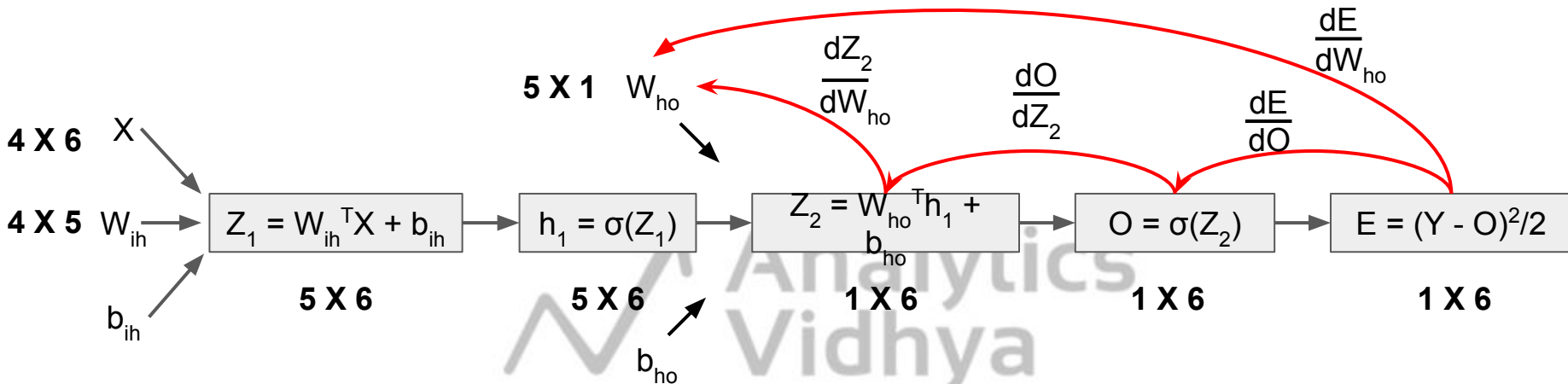


$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}} \quad \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right]^T$$

$$\frac{dE}{dO} * \frac{dO}{dZ_2} = (1 \times 6)$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically

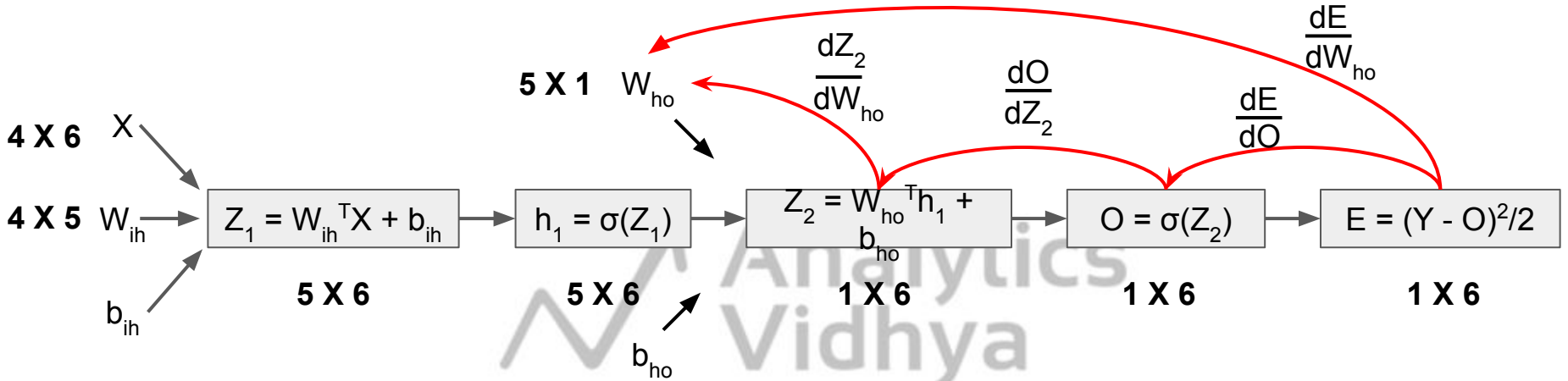


$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}} \quad \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right]^T = (6 \times 1)$$

$$\frac{dE}{dO} * \frac{dO}{dZ_2} = (1 \times 6)$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically

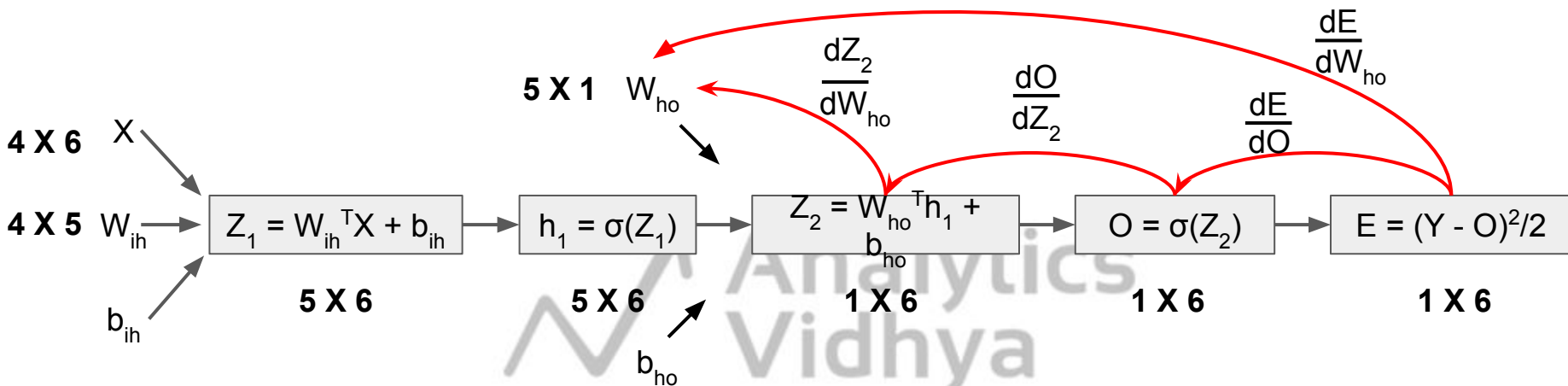


$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}} \quad \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right]^T = (6 \times 1) \quad \frac{dZ_2}{dW_{ho}} \times \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right]^T = (5 \times 6) \times (6 \times 1)$$

$$\frac{dE}{dO} * \frac{dO}{dZ_2} = (1 \times 6)$$

$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically

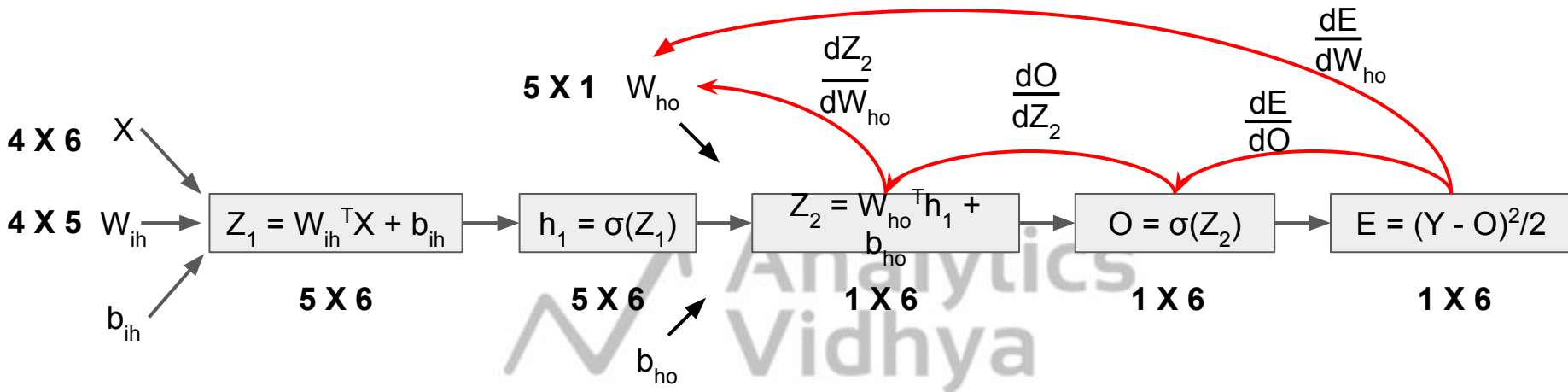


$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}} \quad \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right]^T = (6 \times 1) \quad \frac{dZ_2}{dW_{ho}} \times \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right]^T = (5 \times 1)$$

$$\frac{dE}{dO} * \frac{dO}{dZ_2} = (1 \times 6)$$

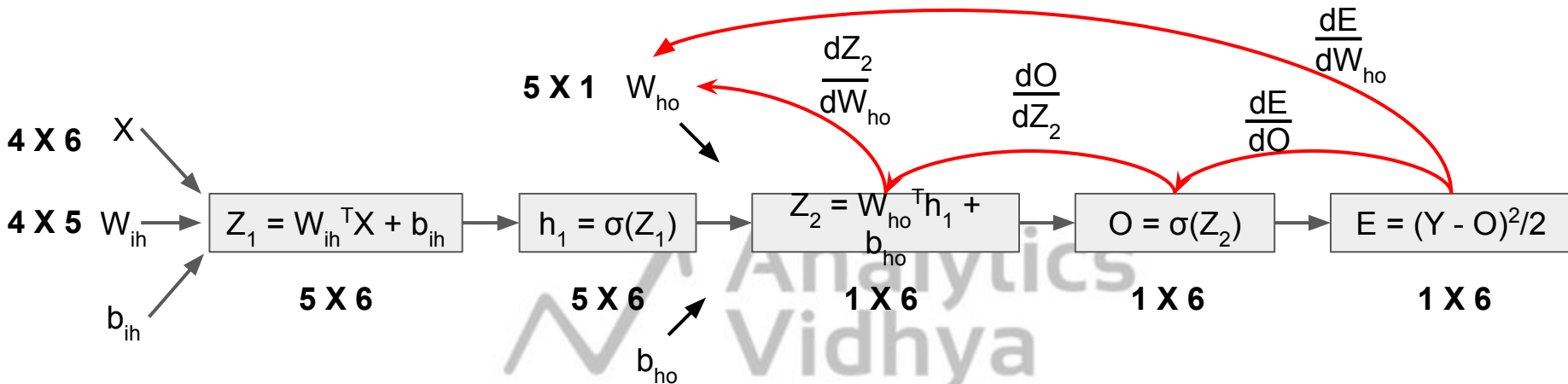
$$\frac{dZ_2}{dW_{ho}} = h_1 \longrightarrow 5 \times 6$$

Understanding Backward Propagation Mathematically



$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$

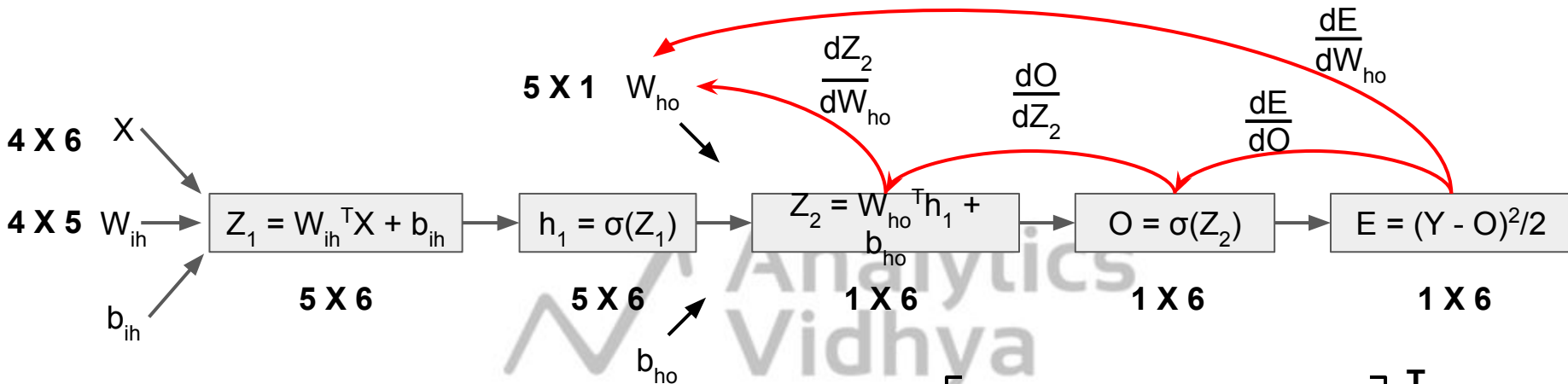
Understanding Backward Propagation Mathematically



~~$$\frac{dE}{dW_{ho}} = \frac{dE}{dO} * \frac{dO}{dZ_2} * \frac{dZ_2}{dW_{ho}}$$~~

$$\frac{dE}{dW_{ho}} = \frac{dZ_2}{dW_{ho}} \times \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right]^T$$

Understanding Backward Propagation Mathematically



$$\frac{dE}{dW_{ho}} = \frac{dZ_2}{dW_{ho}} \times \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right]^T$$

$$\frac{dE}{dW_{ih}} = \frac{dZ_1}{dW_{ih}} \times \left[\frac{dh_1}{dZ_1} * \frac{dZ_2}{dW_{ho}} \times \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right] \right]^T$$

$$\frac{dE}{db_{ho}} = \frac{dZ_2}{db_{ho}} \times \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right]^T$$

$$\frac{dE}{db_{ih}} = \frac{dZ_1}{db_{ih}} \times \left[\frac{dh_1}{dZ_1} * \frac{dZ_2}{dW_{ho}} \times \left[\frac{dE}{dO} * \frac{dO}{dZ_2} \right] \right]^T$$



Thank You