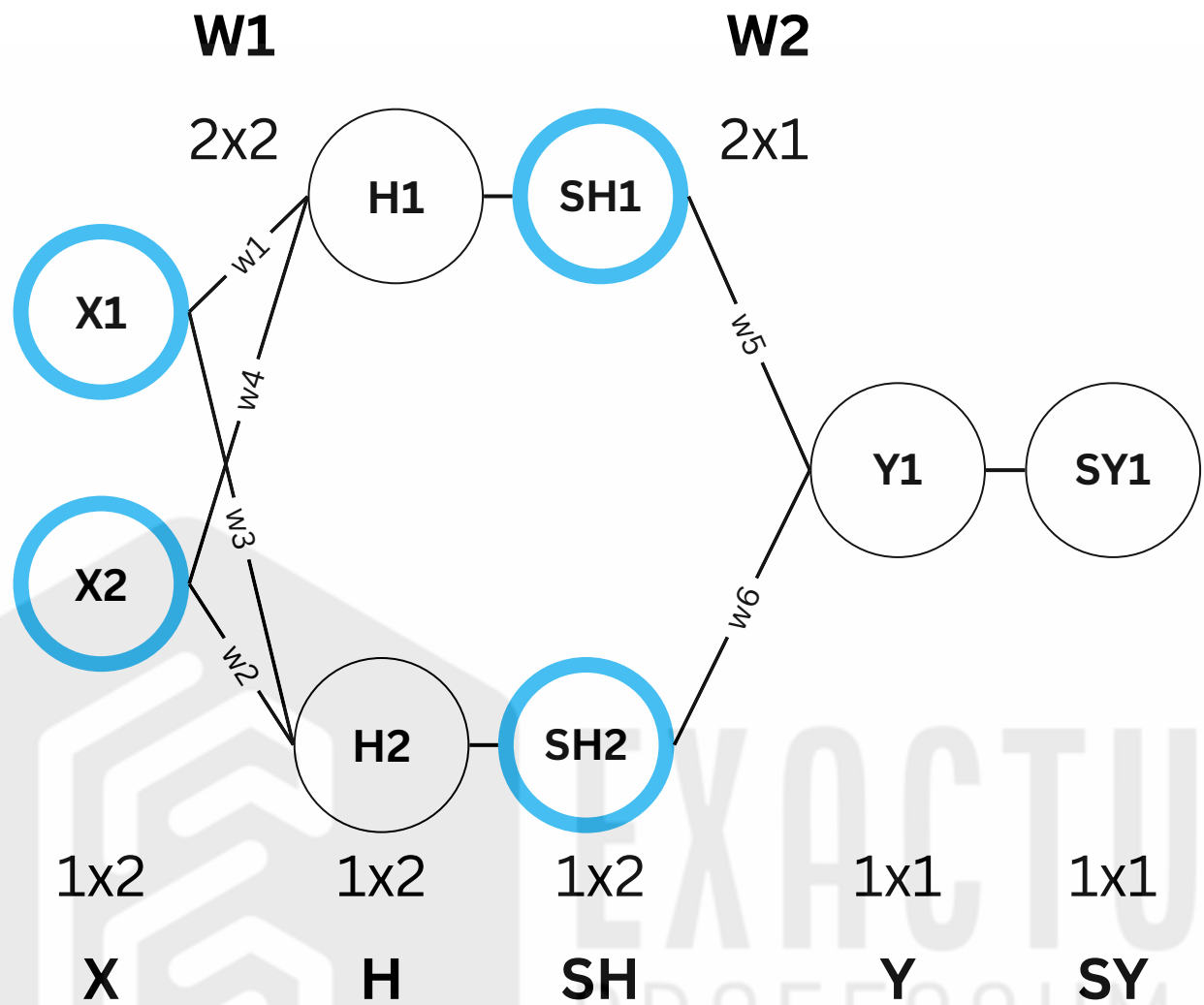


Uma camada escondida



$$\begin{bmatrix} w_5^{new} \\ w_6^{new} \end{bmatrix}_{2 \times 1} = \begin{bmatrix} w_5^{old} \\ w_6^{old} \end{bmatrix}_{2 \times 1} - LR * \begin{bmatrix} \frac{\partial E}{\partial w_5} \\ \frac{\partial E}{\partial w_6} \end{bmatrix}_{2 \times 1}$$

$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial H3}{\partial w_5}$$

$$\frac{\partial E}{\partial w_6} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial w_6}$$

Uma camada escondida

Algebricamente:

$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial w_5} = -(S - SY) * SY * (1 - SY) * SH1$$

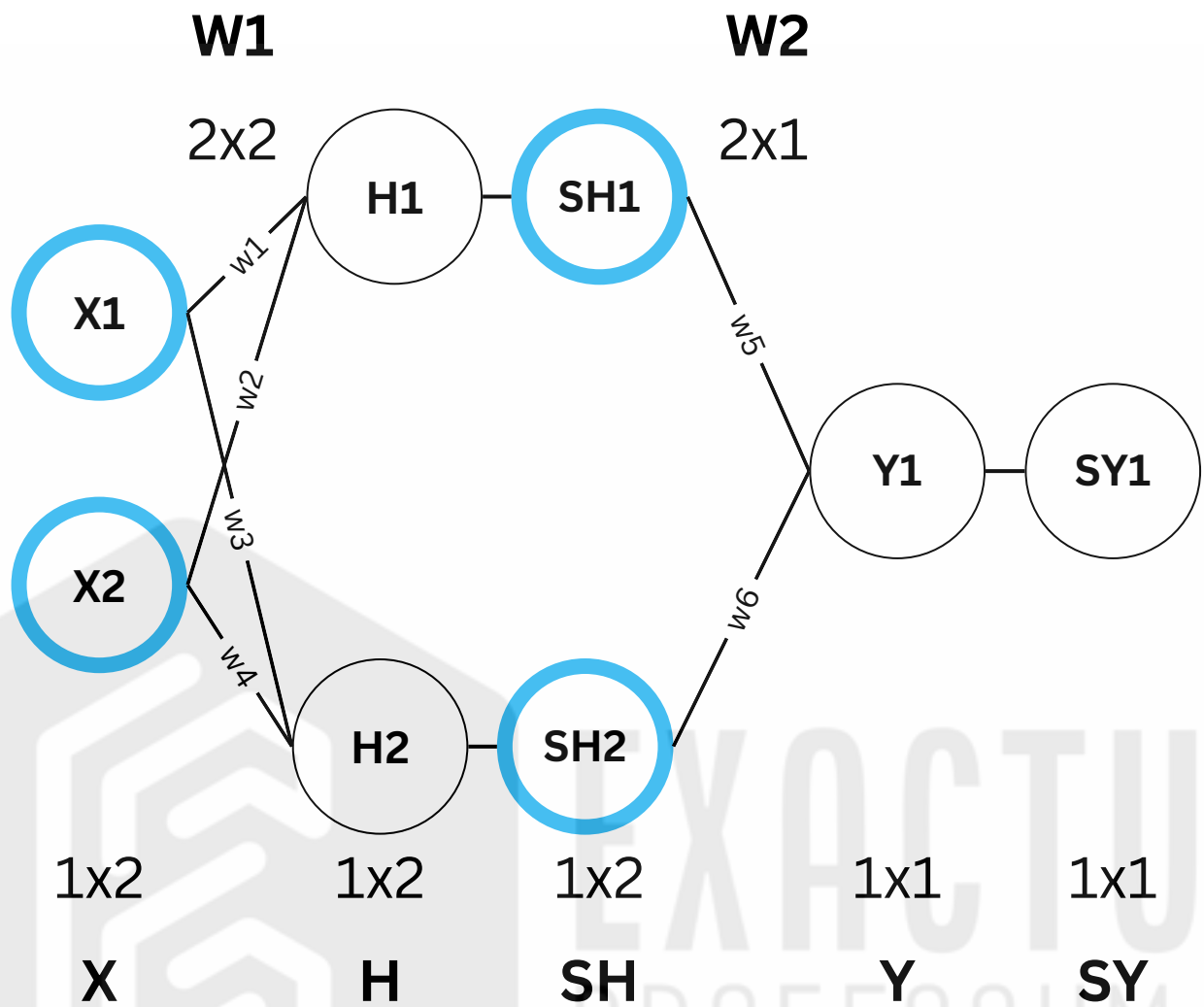
$$\frac{\partial E}{\partial w_6} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial w_6} = -(S - SY) * SY * (1 - SY) * SH2$$

Matricialmente:

$$\begin{bmatrix} \frac{\partial E}{\partial w_{5|6}} \end{bmatrix}_{2 \times 1} = \left(\begin{bmatrix} \frac{\partial Y}{\partial w_{5|6}} \end{bmatrix}_{1 \times 2} \right) \cdot \text{dot} \left(\begin{bmatrix} \frac{\partial E}{\partial SY} \end{bmatrix}_{1 \times 1} * \begin{bmatrix} \frac{\partial SY}{\partial Y} \end{bmatrix}_{1 \times 1} \right) = (SH_{1 \times 2}^T) \cdot \text{dot}([-(S_{1 \times 1} - SY_{1 \times 1})]_{1 \times 1} * [SY_{1 \times 1} * (1 - SY_{1 \times 1})]_{1 \times 1})$$

Correção dos pesos na camada de saída

Uma camada escondida



$$\begin{bmatrix} w_1^{new} & w_3^{new} \\ w_2^{new} & w_4^{new} \end{bmatrix}_{2 \times 2} = \begin{bmatrix} w_1^{old} & w_3^{old} \\ w_2^{old} & w_4^{old} \end{bmatrix}_{2 \times 2} - LR * \begin{bmatrix} \frac{\partial E}{\partial w_1} & \frac{\partial E}{\partial w_3} \\ \frac{\partial E}{\partial w_2} & \frac{\partial E}{\partial w_4} \end{bmatrix}_{2 \times 2}$$

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH1} \frac{\partial SH1}{\partial H1} \frac{\partial H1}{\partial w_1}$$

$$\frac{\partial E}{\partial w_2} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH1} \frac{\partial SH1}{\partial H1} \frac{\partial H1}{\partial w_2}$$

$$\frac{\partial E}{\partial w_3} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH2} \frac{\partial SH2}{\partial H2} \frac{\partial H2}{\partial w_3}$$

$$\frac{\partial E}{\partial w_4} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH2} \frac{\partial SH2}{\partial H2} \frac{\partial H2}{\partial w_4}$$

Uma camada escondida

Algebricamente:

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH1} \frac{\partial SH1}{\partial H1} \frac{\partial H1}{\partial w_1} = -(S - SY) * SY * (1 - SY) * W_5 * SH1 * (1 - SH1) * X_1$$

$$\frac{\partial E}{\partial w_2} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH1} \frac{\partial SH1}{\partial H1} \frac{\partial H1}{\partial w_2} = -(S - SY) * SY * (1 - SY) * W_5 * SH1 * (1 - SH1) * X_2$$

$$\frac{\partial E}{\partial w_3} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH2} \frac{\partial SH2}{\partial H2} \frac{\partial H2}{\partial w_3} = -(S - SY) * SY * (1 - SY) * W_6 * SH2 * (1 - SH2) * X_1$$

$$\frac{\partial E}{\partial w_4} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH2} \frac{\partial SH2}{\partial H2} \frac{\partial H2}{\partial w_4} = -(S - SY) * SY * (1 - SY) * W_6 * SH2 * (1 - SH2) * X_2$$

Matricialmente:

$$\left[\frac{\partial E}{\partial w_{1|2|3|4}} \right]_{2 \times 2} = (X_{1 \times 2}^T) \cdot \text{dot} \left(\left(\left[\frac{\partial E}{\partial SY} \right]_{1 \times 1} * \left[\frac{\partial SY}{\partial Y} \right]_{1 \times 1} \right)_{1 \times 1} \cdot \text{dot} \left((\text{deriv}(SH_{1 \times 2}) * W_{2 \times 1}^T)_{1 \times 2} \right)_{1 \times 2} \right)$$

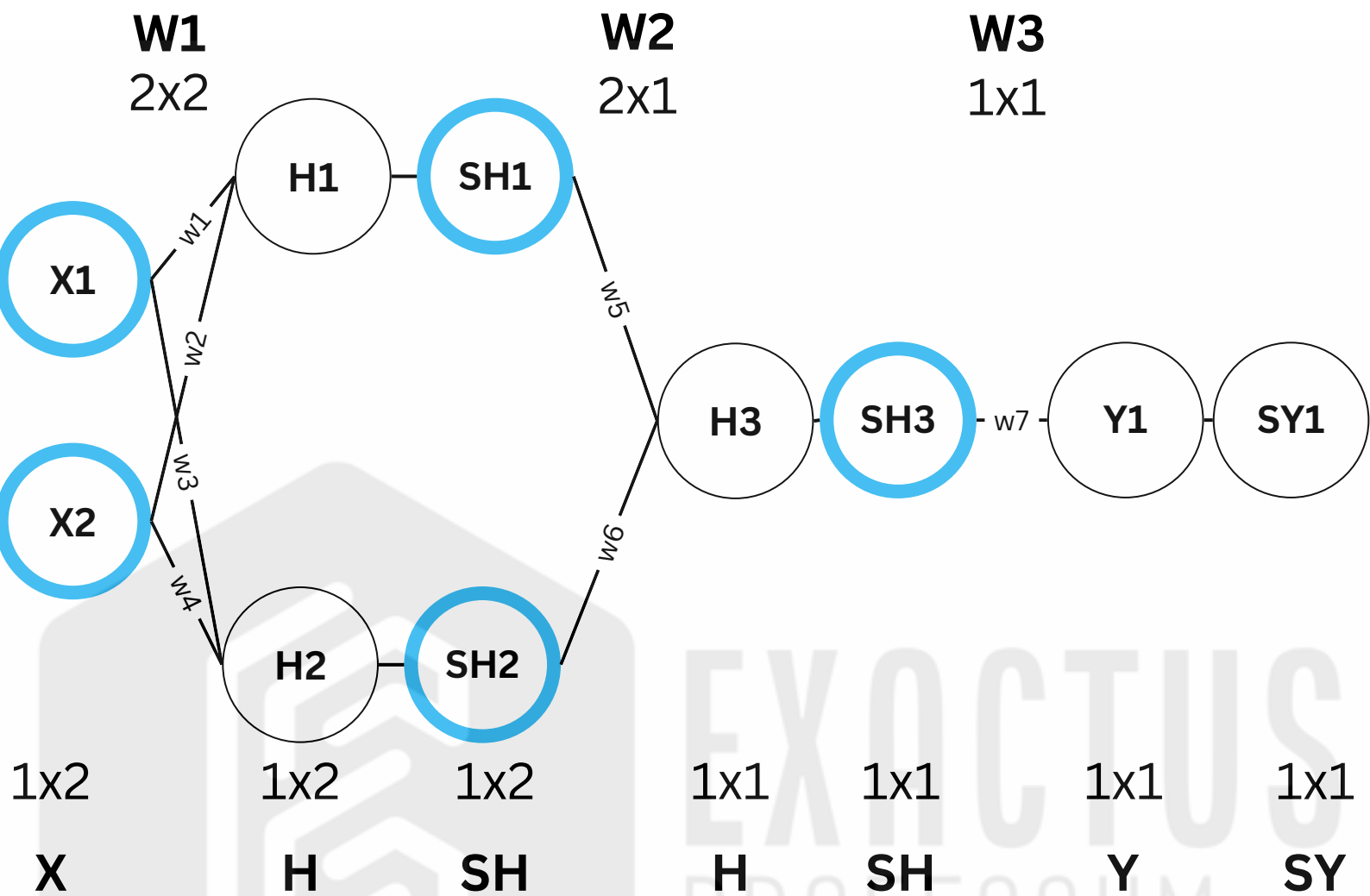
Correção dos pesos na camada de entrada

Uma camada escondida

$$\left[\frac{\partial E}{\partial w_{1|2|3|4}} \right]_{2 \times 2} = (X_{1 \times 2}^T) \cdot \text{dot} \left(\left(\left[\frac{\partial E}{\partial SY} \right]_{1 \times 1} * \left[\frac{\partial SY}{\partial Y} \right]_{1 \times 1} \right) \cdot \text{dot} \left((\text{deriv}(SH_{1 \times 2}) * W_{2 \times 1}^T)_{1 \times 2} \right) \right)$$
$$\left[\frac{\partial E}{\partial w_{5|6}} \right]_{2 \times 1} = (SH_{1 \times 2}^T) \cdot \text{dot} \left(\left(\left[\frac{\partial E}{\partial SY} \right]_{1 \times 1} * \left[\frac{\partial SY}{\partial Y} \right]_{1 \times 1} \right) \right)$$

Resumo dos gradientes

Duas camadas escondidas



$$[w_7^{new}]_{1 \times 1} = [w_7^{old}]_{2 \times 1} - LR * \left[\frac{\partial E}{\partial w_7} \right]_{2 \times 1}$$

$$\frac{\partial E}{\partial w_7} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial w_7}$$

Duas camadas escondidas

Algebricamente:

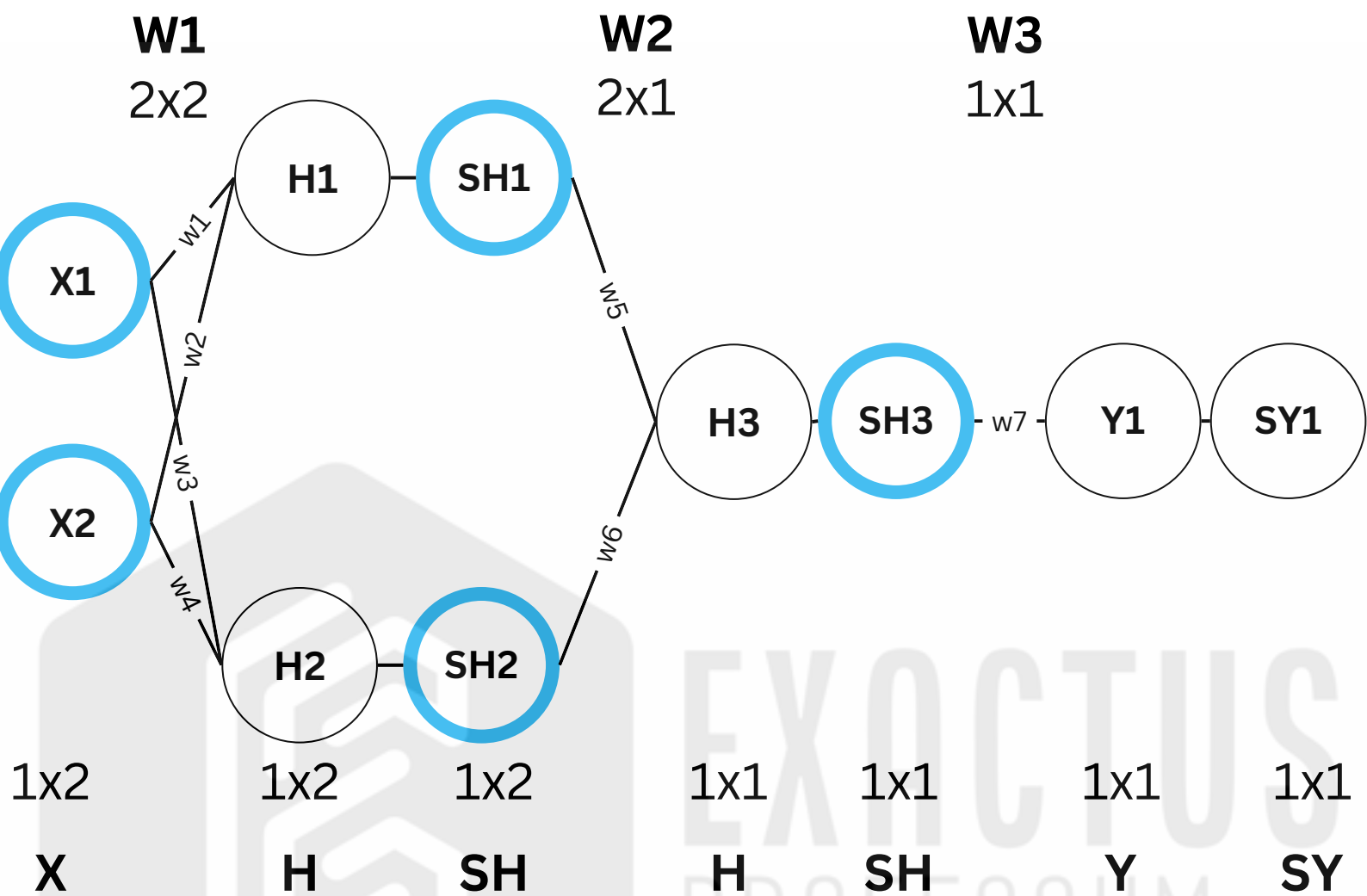
$$\frac{\partial E}{\partial w_7} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial w_7} = -(S - SY) * SY * (1 - SY) * SH3$$

Matricialmente:

$$\left[\frac{\partial E}{\partial w_7} \right]_{1 \times 1} = \left(\left[\frac{\partial Y}{\partial w_7} \right]_{1 \times 1}^T \right) \cdot \text{dot} \left(\left[\frac{\partial E}{\partial SY} \right]_{1 \times 1} * \left[\frac{\partial SY}{\partial Y} \right]_{1 \times 1} \right) = (SH_{1 \times 1}^T) \cdot \text{dot}([- (S_{1 \times 1} - SY_{1 \times 1})]_{1 \times 1} * [SY_{1 \times 1} * (1 - SY_{1 \times 1})]_{1 \times 1})$$

Correção dos pesos na camada de saída

Duas camadas escondidas



$$\begin{bmatrix} w_5^{new} \\ w_6^{new} \end{bmatrix}_{2 \times 1} = \begin{bmatrix} w_5^{old} \\ w_6^{old} \end{bmatrix}_{2 \times 1} - LR * \begin{bmatrix} \frac{\partial E}{\partial w_5} \\ \frac{\partial E}{\partial w_6} \end{bmatrix}_{2 \times 1}$$

$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial w_5}$$

$$\frac{\partial E}{\partial w_6} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial Y}{\partial w_6}$$

Duas camadas escondidas

Algebricamente:

$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial w_5} = [-(S - SY)] * [SY * (1 - SY)] * [w7] * [SH3 * (1 - SH3)] * [SH1]$$

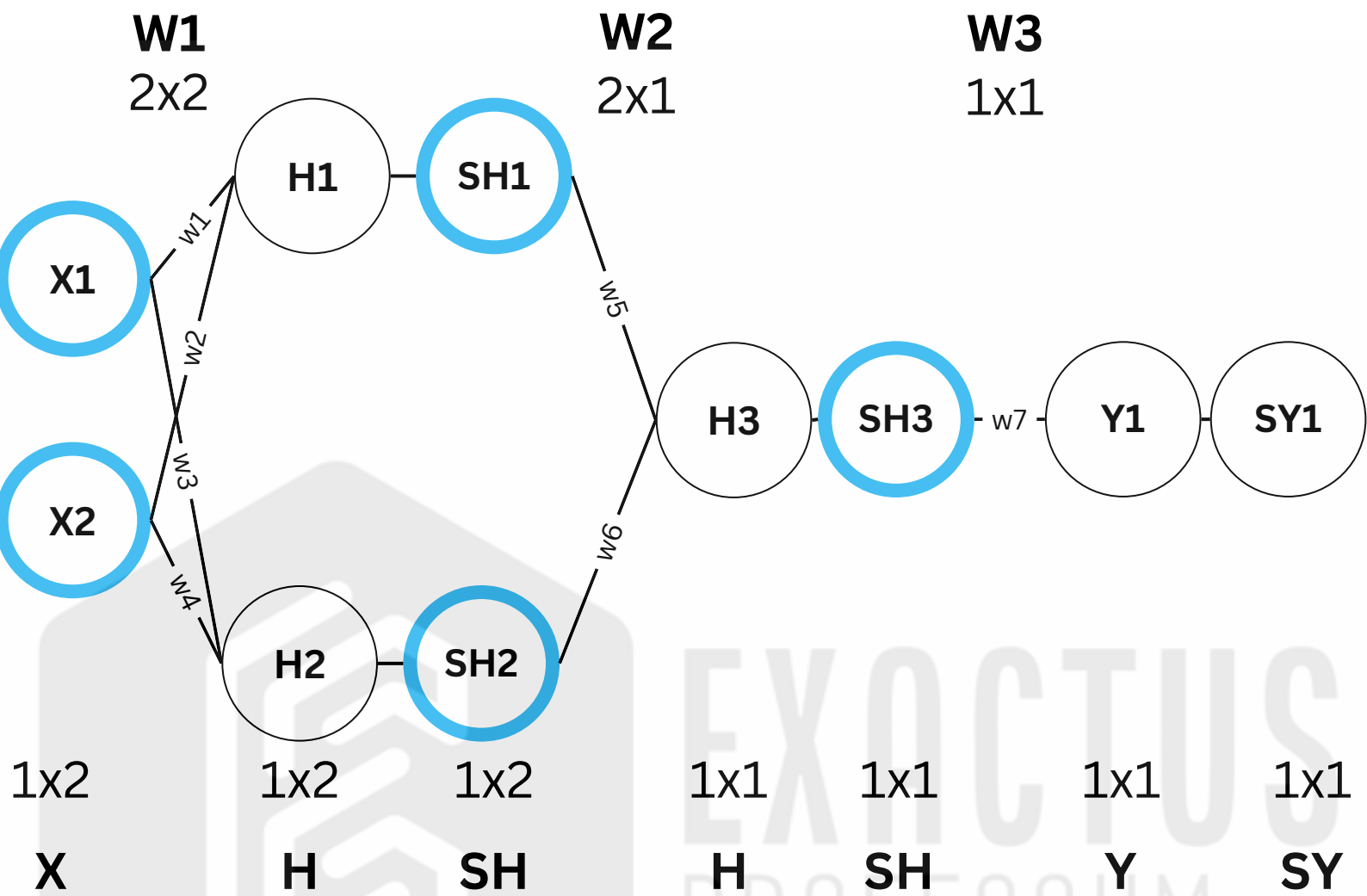
$$\frac{\partial E}{\partial w_6} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial w_6} = [-(S - SY)] * [SY * (1 - SY)] * [w7] * [SH3 * (1 - SH3)] * [SH2]$$

Matricialmente:

$$\begin{aligned} \begin{bmatrix} \frac{\partial E}{\partial w_{5|6}} \end{bmatrix}_{2 \times 1} &= \left(\begin{bmatrix} \frac{\partial Y}{\partial w_{5|6}} \end{bmatrix}_{1 \times 2}^T \right)_{2 \times 1} \cdot \text{dot} \left(\left(\begin{bmatrix} \frac{\partial E}{\partial SY} \end{bmatrix}_{1 \times 1} * \begin{bmatrix} \frac{\partial SY}{\partial Y} \end{bmatrix}_{1 \times 1} \right)_{1 \times 1} \cdot \text{dot} \left(\begin{bmatrix} \frac{\partial Y}{\partial SH3} \end{bmatrix}_{1 \times 1} * \begin{bmatrix} \frac{\partial SH3}{\partial H3} \end{bmatrix}_{1 \times 1} \right)_{1 \times 1} \right) \\ &= (SH_{1 \times 2}^T)_{2 \times 1} \cdot \text{dot}((- (S_{1 \times 1} - SY_{1 \times 1}))_{1 \times 1} * [SY_{1 \times 1} * (1 - SY_{1 \times 1})]_{1 \times 1})_{1 \times 1} \cdot \text{dot}([W_{1 \times 1}] * [SH_{1 \times 1}])_{1 \times 1}) \\ \begin{bmatrix} \frac{\partial E}{\partial w_{5|6}} \end{bmatrix}_{2 \times 1} &= (SH_{1 \times 2}^T)_{2 \times 1} \cdot \text{dot} \left(\left(\begin{bmatrix} \frac{\partial E}{\partial SY} \end{bmatrix}_{1 \times 1} * \begin{bmatrix} \frac{\partial SY}{\partial Y} \end{bmatrix}_{1 \times 1} \right)_{1 \times 1} \cdot \text{dot} \left(\begin{bmatrix} \frac{\partial Y}{\partial SH3} \end{bmatrix}_{1 \times 1} * \begin{bmatrix} \frac{\partial SH3}{\partial H3} \end{bmatrix}_{1 \times 1} \right)_{1 \times 1} \right) \end{aligned}$$

Correção dos pesos na camada escondida

Duas camadas escondidas



$$\begin{bmatrix} w_1^{new} & w_3^{new} \\ w_2^{new} & w_4^{new} \end{bmatrix}_{2 \times 2} = \begin{bmatrix} w_1^{old} & w_3^{old} \\ w_2^{old} & w_4^{old} \end{bmatrix}_{2 \times 2} - LR * \begin{bmatrix} \frac{\partial E}{\partial w_1} & \frac{\partial E}{\partial w_3} \\ \frac{\partial E}{\partial w_2} & \frac{\partial E}{\partial w_4} \end{bmatrix}_{2 \times 2}$$

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial SH1} \frac{\partial SH1}{\partial H1} \frac{\partial H1}{\partial w_1} = -(S - SY) * SY * (1 - SY) * W_7 * SH3 * (1 - SH3) * W_5 * SH1 * (1 - SH1) * X_1$$

$$\frac{\partial E}{\partial w_2} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial SH1} \frac{\partial SH1}{\partial H1} \frac{\partial H1}{\partial w_2} = -(S - SY) * SY * (1 - SY) * W_7 * SH3 * (1 - SH3) * W_5 * SH1 * (1 - SH1) * X_2$$

$$\frac{\partial E}{\partial w_3} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial SH2} \frac{\partial SH2}{\partial H2} \frac{\partial H2}{\partial w_3} = -(S - SY) * SY * (1 - SY) * W_7 * SH3 * (1 - SH3) * W_6 * SH1 * (1 - SH1) * X_1$$

$$\frac{\partial E}{\partial w_4} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial SH2} \frac{\partial SH2}{\partial H2} \frac{\partial H2}{\partial w_4} = -(S - SY) * SY * (1 - SY) * W_7 * SH3 * (1 - SH3) * W_6 * SH1 * (1 - SH1) * X_2$$

Duas camadas escondidas

Algebricamente:

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial SH1} \frac{\partial SH1}{\partial H1} \frac{\partial H1}{\partial w_1} = -(S - SY) * SY * (1 - SY) * W_7 * SH3 * (1 - SH3) * W_5 * SH1 * (1 - SH1) * X_1$$

$$\frac{\partial E}{\partial w_2} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial SH1} \frac{\partial SH1}{\partial H1} \frac{\partial H1}{\partial w_2} = -(S - SY) * SY * (1 - SY) * W_7 * SH3 * (1 - SH3) * W_5 * SH1 * (1 - SH1) * X_2$$

$$\frac{\partial E}{\partial w_3} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial SH2} \frac{\partial SH2}{\partial H2} \frac{\partial H2}{\partial w_3} = -(S - SY) * SY * (1 - SY) * W_7 * SH3 * (1 - SH3) * W_6 * SH1 * (1 - SH1) * X_1$$

$$\frac{\partial E}{\partial w_4} = \frac{\partial E}{\partial SY} \frac{\partial SY}{\partial Y} \frac{\partial Y}{\partial SH3} \frac{\partial SH3}{\partial H3} \frac{\partial H3}{\partial SH2} \frac{\partial SH2}{\partial H2} \frac{\partial H2}{\partial w_4} = -(S - SY) * SY * (1 - SY) * W_7 * SH3 * (1 - SH3) * W_6 * SH1 * (1 - SH1) * X_2$$

Matricialmente:

$$\begin{bmatrix} \frac{\partial E}{\partial w_{1|2|3|4}} \end{bmatrix}_{2 \times 2} = \left(\begin{bmatrix} \frac{\partial H}{\partial w_{1|2|3|4}} \end{bmatrix}_{1 \times 2}^T \right) \cdot \text{dot} \left(\left(\begin{bmatrix} \frac{\partial E}{\partial SY} \end{bmatrix}_{1 \times 1} * \begin{bmatrix} \frac{\partial SY}{\partial Y} \end{bmatrix}_{1 \times 1} \right) \cdot \text{dot} \left(\begin{bmatrix} \frac{\partial Y}{\partial SH} \end{bmatrix}_{1 \times 1} * \begin{bmatrix} \frac{\partial SH}{\partial H} \end{bmatrix}_{1 \times 1} \right) \cdot \text{dot} \left(\begin{bmatrix} \frac{\partial H}{\partial SH} \end{bmatrix}_{2 \times 1}^T * \begin{bmatrix} \frac{\partial SH}{\partial H} \end{bmatrix}_{1 \times 2} \right) \right)$$

$$\begin{bmatrix} \frac{\partial E}{\partial w_{1|2|3|4}} \end{bmatrix}_{2 \times 2} = (X_{1 \times 2}^T)_{2 \times 1} \cdot \text{dot} \left(\left(\begin{bmatrix} \frac{\partial E}{\partial SY} \end{bmatrix}_{1 \times 1} * \begin{bmatrix} \frac{\partial SY}{\partial Y} \end{bmatrix}_{1 \times 1} \right) \cdot \text{dot} ((W_{1 \times 1}^T * \text{deriv}(SH_{1 \times 1}))_{1 \times 1})_{1 \times 1} \cdot \text{dot} ((W_{2 \times 1}^T * \text{deriv}(SH_{1 \times 2}))_{1 \times 2})_{1 \times 2} \right)$$

Correção dos pesos na camada de entrada

Duas camadas escondidas

$$\left[\frac{\partial E}{\partial w_7} \right]_{1 \times 1} = (SH_{1 \times 1}^T)_{1 \times 1} \cdot \text{dot} \left(\underbrace{\left[\frac{\partial E}{\partial SY} \right]_{1 \times 1} * \left[\frac{\partial SY}{\partial Y} \right]_{1 \times 1}}_{\text{Derivadas Acumuladas}} \right)$$

$$\left[\frac{\partial E}{\partial w_{5|6}} \right]_{2 \times 1} = (SH_{1 \times 2}^T)_{2 \times 1} \cdot \text{dot} \left(\underbrace{\left(\underbrace{\left[\frac{\partial E}{\partial SY} \right]_{1 \times 1} * \left[\frac{\partial SY}{\partial Y} \right]_{1 \times 1}}_{\text{Derivada Anterior}} \right)_{1 \times 1} \cdot \text{dot} \left(\left[\frac{\partial Y}{\partial SH} \right]_{1 \times 1}^T * \left[\frac{\partial SH}{\partial H} \right]_{1 \times 1} \right)_{1 \times 1}}_{\text{Derivadas Acumuladas}} \right)$$

$$\left[\frac{\partial E}{\partial w_{1|2|3|4}} \right]_{2 \times 2} = (X_{1 \times 2}^T)_{2 \times 1} \cdot \text{dot} \left(\underbrace{\left(\left(\left(\left[\frac{\partial E}{\partial SY} \right]_{1 \times 1} * \left[\frac{\partial SY}{\partial Y} \right]_{1 \times 1} \right)_{1 \times 1} \cdot \text{dot} \left(\left[\frac{\partial Y}{\partial SH} \right]_{1 \times 1}^T * \left[\frac{\partial SH}{\partial H} \right]_{1 \times 1} \right)_{1 \times 1} \cdot \text{dot} \left(\left[\frac{\partial H}{\partial SH} \right]_{2 \times 1}^T * \left[\frac{\partial SH}{\partial H} \right]_{1 \times 2} \right)_{1 \times 2} \right)_{1 \times 1}}_{\text{Derivada Anterior}} \right)_{1 \times 1}}_{\text{Derivadas Acumuladas}} \right)$$

Resumo dos gradientes

Duas camadas escondidas

Derivadas Acumuladas lista

Índice 0

$$(X_{1 \times 2}^T)_{2 \times 1} \cdot \text{dot} \left(\left[\frac{\partial E}{\partial SY} \right]_{1 \times 1} * \left[\frac{\partial SY}{\partial Y} \right]_{1 \times 1} \right)_{1 \times 1} \cdot \text{dot} \left(\left[\frac{\partial Y}{\partial SH} \right]_{1 \times 1}^T * \left[\frac{\partial SH}{\partial H} \right]_{1 \times 1} \right)_{1 \times 1} \cdot \text{dot} \left(\left[\frac{\partial H}{\partial SH} \right]_{2 \times 1}^T * \left[\frac{\partial SH}{\partial H} \right]_{1 \times 2} \right)_{1 \times 2}$$

Corrige os pesos da camada de **entrada**

Índice 1

$$(SH_{1 \times 2}^T)_{2 \times 1} \cdot \text{dot} \left(\left[\frac{\partial E}{\partial SY} \right]_{1 \times 1} * \left[\frac{\partial SY}{\partial Y} \right]_{1 \times 1} \right)_{1 \times 1} \cdot \text{dot} \left(\left[\frac{\partial Y}{\partial SH} \right]_{1 \times 1}^T * \left[\frac{\partial SH}{\partial H} \right]_{1 \times 1} \right)_{1 \times 1}$$

Corrige os pesos da camada **escondida**

Índice 2

$$(SH_{1 \times 1}^T)_{1 \times 1} \cdot \text{dot} \left(\left[\frac{\partial E}{\partial SY} \right]_{1 \times 1} * \left[\frac{\partial SY}{\partial Y} \right]_{1 \times 1} \right)_{1 \times 1}$$

Corrige os pesos da camada de **saída**