Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Industrial

Tarea 2

Integrantes:	Michael Clemans
	Paula Marín
	Bárbara Pérez
	Sebastián Urbina
Profesor:	Ángel Jiménez Molina
Auxiliar:	Rafael De la Sotta Vargas

Fecha de entrega: 14 de noviembre de 2021

Pregunta 1

1.1 Forward propagation

A continuación se puede observar la arquitectura de la red feedforward. Ésta posee 2 capas. Con la función sigmoid($\sigma()$ de activación a la salida de cada una.

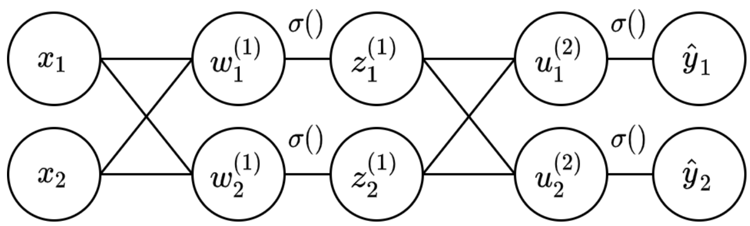

Figura 1: Arquitectura red feedforward

Para proceder con el algoritmo *forward propagation* se identifican los siguientes datos a partir del enunciado. Los cuales se utilizarán matricialmente.

$$X = \begin{pmatrix} 0.05 \\ 0.1 \end{pmatrix} \quad Y = \begin{pmatrix} 0.01 \\ 0.99 \end{pmatrix}$$

Donde X representa el vector de entrada e Y contiene los valores objetivo.

$$W^{(1)} = \begin{pmatrix} 0.15 & 0.2 \\ 0.25 & 0.3 \end{pmatrix} \quad U = \begin{pmatrix} 0.4 & 0.45 \\ 0.5 & 0.55 \end{pmatrix}$$

$$w_0^{(1)} = \begin{pmatrix} 0.35 \\ 0.35 \end{pmatrix} \quad w_s^{(2)} = \begin{pmatrix} 0.6 \\ 0.6 \end{pmatrix}$$

Aquí, $w_0^{(1)}$ contiene los pesos entre la capa de entrada y la capa oculta. U contiene los pesos entre la capa oculta y la salida. $w_0^{(1)}, w_s^{(1)}$ representan los *bias* de ambas capas.

Comenzamos con el algoritmo de *forward propagation* computando el output de la capa oculta. El cual se obtiene luego de ponderar los pesos correspondiente con los datos de entrada y finalmente aplicarle la función *sigmoid*.

$$
\begin{aligned}
Z^{(1)} &= \begin{pmatrix} z_1^{(1)} \\ z_2^{(1)} \end{pmatrix} \\
&= \sigma(W^{(1)} \cdot X + w_0^{(1)}) \\
&= \sigma\left(\begin{pmatrix} 0.15 & 0.2 \\ 0.25 & 0.3 \end{pmatrix} \begin{pmatrix} 0.05 \\ 0.1 \end{pmatrix} + \begin{pmatrix} 0.35 \\ 0.35 \end{pmatrix} \right) \\
&= \sigma\left(\begin{pmatrix} 0.028 \\ 0.043 \end{pmatrix} + \begin{pmatrix} 0.35 \\ 0.35 \end{pmatrix} \right) \\
&= \sigma\left(\begin{pmatrix} 0.378 \\ 0.393 \end{pmatrix} \right) \\
&= \begin{pmatrix} \frac{1}{1+e^{-0.378}} \\ \frac{1}{1+e^{-0.393}} \end{pmatrix} \\
&= \begin{pmatrix} 0.593 \\ 0.597 \end{pmatrix}
\end{aligned}
$$

El segundo paso hacia delante ocurre desde la capa oculta hasta la de salida, siguiendo la misma lógica anterior se llega a los valores estimados $\hat{Y}$ de la salida.

$$
\begin{aligned}
\hat{Y} &= \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \end{pmatrix} \\
&= \sigma(U \cdot Z^{(1)} + w_s^{(2)}) \\
&= \sigma\left(\begin{pmatrix} 0.4 & 0.45 \\ 0.5 & 0.55 \end{pmatrix} \begin{pmatrix} 0.593 \\ 0.597 \end{pmatrix} + \begin{pmatrix} 0.6 \\ 0.6 \end{pmatrix} \right) \\
&= \sigma\left(\begin{pmatrix} 0.506 \\ 0.625 \end{pmatrix} + \begin{pmatrix} 0.6 \\ 0.6 \end{pmatrix} \right) \\
&= \sigma\left(\begin{pmatrix} 1.106 \\ 1.225 \end{pmatrix} \right) \\
&= \begin{pmatrix} \frac{1}{1+e^{1.106}} \\ \frac{1}{1+e^{1.225}} \end{pmatrix} \\
&= \begin{pmatrix} 0.751 \\ 0.772 \end{pmatrix}
\end{aligned}
$$

Finalmente, calculamos la pérdida utilizando la fórmula del error cuadrático medio.

$$
\begin{aligned}
L &= \frac{1}{2}||Y - \hat{Y}||_2^2 \\
&= \frac{1}{2}[(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2] \\
&= \frac{1}{2}[(0.01 - 0.751)^2 + (0.99 - 0.772)^2] \\
&= 0.2983
\end{aligned}
$$

1.2 Backward propagation

Utilizando el algoritmo de backpropagation se computó manualmente el gradiente de L como se muestra a continuación:

1.2.1 Pesos capa salida

$$\frac{\partial L}{\partial u_{1,1}} = \frac{\partial L}{\partial \hat{y}_1} \cdot \frac{\partial \hat{y}_1}{\partial u_1^{(2)}} \cdot \frac{\partial u_1^{(2)}}{\partial u_{1,1}}$$

Vamos a proceder a obtener cada termino de manera particular.

$$\begin{aligned}\frac{\partial L}{\partial \hat{y}_1} &= \frac{\partial}{\partial \hat{y}_1}\left(\frac{1}{2}[(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2]\right)\\ &= \hat{y}_1 - y_1\\ &= 0.751 - 0.01\\ &= 0.741\end{aligned}$$

Notemos que $\hat{y}_i = \sigma(u_i^{(2)}) = \frac{1}{1+e^{-u_i^{(2)}}}, \quad i \in \{1, 2\}$. Además, $\frac{\partial \sigma(x)}{\partial x} = \sigma(x)(1 - \sigma(x))$

$$\frac{\partial \hat{y}_1}{\partial u_1^{(2)}} = \sigma(u_2^{(2)})(1 - \sigma(u_2^{(2)})) = \hat{y}_1(1 - \hat{y}_1) = 0.751(1 - 0.751) = 0.187$$

Luego, tenemos que:

$$u_1^{(2)} = z_1^{(1)} u_{1,1} + z_2^{(1)} u_{1,2}$$

Donde, de los resultados anteriores, $z_1^{(1)} = 0.593, z_2^{(1)} = 0.597, u_{1,1} = 0.4, u_{1,2} = 0.45$

Por lo que,

$$\frac{\partial u_1^{(1)}}{\partial u_{1,1}} = z_1^{(1)} = 0.593$$

Finalmente podemos calcular:

$$\frac{\partial L}{\partial u_{1,1}} = \frac{\partial L}{\partial \hat{y}_1} \cdot \frac{\partial \hat{y}_1}{\partial u_1^{(2)}} \cdot \frac{\partial u_1}{\partial u_{1,1}} = 0.741 \cdot 0.187 \cdot 0.593 = 0.08217$$

Hacemos lo mismo con respecto a $u_{1,2}$

$$\frac{\partial L}{\partial u_{1,2}} = \frac{\partial L}{\partial \hat{y}_1} \cdot \frac{\partial \hat{y}_1}{\partial u_1^{(2)}} \cdot \frac{\partial u_1^{(2)}}{\partial u_{1,2}}$$

Para este caso sólo nos falta calcular $\frac{\partial u_1}{\partial u_{1,2}}$. De la parte anterior, podemos obtener directamente:

$$\frac{\partial u_1}{\partial u_{1,2}} = z_2^{(1)} = 0.597$$

Por lo tanto,

$$\frac{\partial L}{\partial u_{1,2}} = 0.741 \cdot 0.187 \cdot 0.597 = 0.08272$$

Ahora, debemos seguir el otro camino para obtener las derivadas respecto a $u_{2,1}, u_{2,2}$. Siguiendo el mismo procedimiento de antes, por regla de la cadena, se tiene:

$$\frac{\partial L}{\partial u_{2,1}} = \frac{\partial L}{\partial \hat{y}_2} \cdot \frac{\partial \hat{y}_2}{\partial u_2^{(2)}} \cdot \frac{\partial u_2^{(2)}}{\partial u_{2,1}}$$

Debemos calcular $\frac{\partial L}{\partial \hat{y}_2}, \frac{\partial \hat{y}_2}{\partial u_2^{(2)}}, \frac{\partial u_2^{(2)}}{\partial u_{2,1}}$.

$$\begin{aligned}\frac{\partial L}{\partial \hat{y}_2} &= \frac{\partial}{\partial \hat{y}_2}\left(\frac{1}{2}[(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2]\right) \\ &= \hat{y}_2 - y_2 \\ &= 0.772 - 0.99 \\ &= -0.218\end{aligned}$$

Igual que antes, podemos obtener directamente

$$\frac{\partial \hat{y}_2}{\partial u_2^{(2)}} = \sigma(u_2^{(2)})(1 - \sigma(u_2^{(2)})) = \hat{y}_2(1 - \hat{y}_2) = 0.772(1 - 0.772) = 0.176$$

Para obtener la derivada respecto a $u_2^{(2)}$. Sabemos que:

$$u_2^{(2)} = z_1^{(1)} u_{2,1} + z_2^{(1)} u_{2,2}$$

Por lo que:

$$\frac{\partial u_2}{\partial u_{2,1}} = z_1^{(1)} = 0.593$$

Ya tenemos todo para calcular:

$$\frac{\partial L}{\partial u_{2,1}} = \frac{\partial L}{\partial \hat{y}_2} \cdot \frac{\partial \hat{y}_2}{\partial u_2^{(2)}} \cdot \frac{\partial u_2^{(2)}}{\partial u_{2,1}} = -0.218 \cdot 0.176 \cdot 0.593 = -0.02275$$

Ahora nos falta obtener la derivada de la función de pérdida respecto al peso $u_{2,2}$. Este se puede obtener directamente, porque respecto al último cálculo sólo cambia $\frac{\partial u_2^{(2)}}{\partial u_{2,2}} = z_2^{(1)} = 0.597$.

$$\frac{\partial L}{\partial u_{2,2}} = \frac{\partial L}{\partial \hat{y}_2} \cdot \frac{\partial \hat{y}_2}{\partial u_2^{(2)}} \cdot \frac{\partial u_2^{(2)}}{\partial u_{2,2}} = -0.218 \cdot 0.176 \cdot 0.597 = -0.0229$$

1.2.2 Pesos capa oculta

Aplicando las derivadas parciales a través de lared para llegar al peso $w_{1,1}$ se tiene:

$$\frac{\partial L}{\partial w_{1,1}} = \frac{\partial L}{\partial z_1^{(1)}} \cdot \frac{\partial z_1^{(1)}}{\partial w_1^{(1)}} \cdot \frac{\partial w_1^{(1)}}{\partial w_{1,1}}$$

De donde,

$$\frac{\partial L}{\partial z_1^{(1)}} = \frac{\partial L_{\hat{y}_1}}{\partial z_1^{(1)}} + \frac{\partial L_{\hat{y}_2}}{\partial z_1^{(1)}}$$

En donde empezando por $\frac{\partial L_{\hat{y}_1}}{\partial z_1^{(1)}}$, se tiene:

$$\frac{\partial L_{\hat{y}_1}}{\partial z_1^{(1)}} = \frac{\partial L_{\hat{y}_1}}{\partial u_1^{(1)}} \cdot \frac{\partial u_1^{(1)}}{\partial z_1^{(1)}}$$

En donde se puede calcular $\frac{\partial L_{\hat{y}_1}}{\partial u_1^{(1)}}$ usando valores calculados anteriormente:

$$\begin{aligned}\frac{\partial L_{\hat{y}_1}}{\partial u_1^{(1)}} &= \frac{\partial L}{\partial \hat{y}_1} \cdot \frac{\partial \hat{y}_1}{\partial u_1^{(2)}} \\ &= 0.741 \cdot 0.187 \\ &= 0.138\end{aligned}$$

Y luego $\frac{\partial u_1^{(1)}}{\partial z_1^{(1)}}$ es igual a $u_{1,1}$:

$$u_1^{(2)} = u_{1,1} \cdot z_1 + u_{1,2} \cdot z_2 + 0.6$$

$$\frac{\partial u_1^{(2)}}{\partial z_1^{(1)}} = u_{1,1} = 0.4$$

Con esto se tiene:

$$\begin{aligned}\frac{\partial L_{\hat{y}_1}}{\partial z_1^{(1)}} &= \frac{\partial L_{\hat{y}_1}}{\partial u_1^{(2)}} \cdot \frac{\partial u_2^{(1)}}{\partial z_1^{(1)}} \\ &= 0.138 \cdot 0.40 \\ &= 0.055\end{aligned}$$

Luego siguiendo el mismo proceso para $\frac{\partial L_{\hat{y}_2}}{\partial z_1^{(1)}}$, se obtiene:

$$\frac{\partial L_{\hat{y}_2}}{\partial z_1^{(1)}} = -0.0190$$

Por lo que:

$$\begin{aligned}\frac{\partial L}{\partial z_1^{(1)}} &= \frac{\partial L_{\hat{y}_1}}{\partial z_1^{(1)}} + \frac{\partial L_{\hat{y}_2}}{\partial z_1^{(1)}} \\ &= 0.055 + -0.0190 \\ &= 0.036\end{aligned}$$

Ahora que tenemos $\frac{\partial L}{\partial z_1^{(1)}}$, tenemos que averiguar $\frac{\partial z_1^{(1)}}{\partial w_1^{(1)}}$ y luego $\frac{\partial w_1^{(1)}}{\partial w_{1,1}}$:
Sabemos que:

$$z_1^{(1)} = \frac{1}{1 + e^{-w_1^{(1)}}}$$

$$\begin{aligned}\frac{\partial z_1^{(1)}}{\partial w_1^{(1)}} &= \sigma(w_1^{(1)})(1 - \sigma(w_1^{(1)})) \\ &= z_1^{(1)}(1 - z_1^{(1)}) \\ &= 0.593(1 - 0.593) \\ &= 0.241\end{aligned}$$

Calculamos $\frac{\partial w_1^{(1)}}{\partial w_{1,1}}$. Donde se sabe que $w_1^{(1)}$ se obtiene como:

$$w_1^{(1)} = x_1 w_{1,1} + x_2 w_{1,2}$$

$$\frac{\partial w_1^{(1)}}{\partial w_{1,1}} = x_1 = 0.05$$

Juntando todo:

$$\begin{aligned}\frac{\partial L}{\partial w_{1,1}} &= \frac{\partial L}{\partial z_1^{(1)}} \cdot \frac{\partial z_1^{(1)}}{\partial w_1^{(1)}} \cdot \frac{\partial w_1^{(1)}}{\partial w_{1,1}} \\ &= 0.0363 \cdot 0.241 \cdot 0.05 \\ &= 0.00043\end{aligned}$$

Se repite el paso anterior para $\frac{\partial L}{\partial w_{1,2}}$:

$$\frac{\partial L}{\partial w_{1,2}} = \frac{\partial L}{\partial z_1^{(1)}} \cdot \frac{\partial z_1^{(1)}}{\partial w_1^{(1)}} \cdot \frac{\partial w_1^{(1)}}{\partial w_{1,2}}$$

En donde sólo falta calcular $\frac{\partial w_1^{(1)}}{\partial w_{1,2}}$:

$$\frac{\partial w_1^{(1)}}{\partial w_{1,2}} = x_2 = 0.1$$

Entonces:

$$\begin{aligned}\frac{\partial L}{\partial w_{1,2}} &= 0.0363 \cdot 0.241 \cdot 0.1 \\ &= 0.00087\end{aligned}$$

Se repite el paso anterior para $\frac{\partial L}{\partial w_{2,1}}$:

$$\frac{\partial L}{\partial w_{2,1}} = \frac{\partial L}{\partial z_2^{(1)}} \cdot \frac{\partial z_2^{(1)}}{\partial w_2^{(1)}} \cdot \frac{\partial w_2^{(1)}}{\partial w_{2,1}}$$

En donde se tiene que $\frac{\partial w_2^{(1)}}{\partial w_{2,1}}$ se calcula de:

$$w_2^{(1)} = x_1 w_{2,1} + x_2 w_{2,2}$$

$$\frac{\partial w_2^{(1)}}{\partial w_{2,1}} = x_1 = 0.05$$

además:

$$\begin{aligned}\frac{\partial z_2^{(1)}}{\partial w_2^{(1)}} &= \sigma(w_2^{(1)})(1 - \sigma(w_2^{(1)})) \\ &= z_2^{(1)}(1 - z_2^{(1)}) \\ &= 0.597(1 - 0.597) \\ &= 0.240\end{aligned}$$

Luego se tiene que $\frac{\partial L}{\partial z_2^{(1)}}$ se expresa como:

$$\begin{aligned}\frac{\partial L}{\partial z_2^{(1)}} &= \frac{\partial L_{\hat{y}_1}}{\partial z_2^{(1)}} + \frac{\partial L_{\hat{y}_2}}{\partial z_2^{(1)}} \\ &= \frac{\partial L_{\hat{y}_1}}{\partial u_1^{(2)}} \cdot \frac{\partial u_1^{(2)}}{\partial z_2^{(1)}} + \frac{\partial L_{\hat{y}_2}}{\partial u_2^{(2)}} \cdot \frac{\partial u_2^{(2)}}{\partial z_2^{(1)}} \\ &= \frac{\partial L}{\partial \hat{y}_1} \cdot \frac{\partial \hat{y}_1}{\partial u_1^{(2)}} \cdot \frac{\partial u_1^{(2)}}{\partial z_2^{(1)}} + \frac{\partial L}{\partial \hat{y}_2} \cdot \frac{\partial \hat{y}_2}{\partial u_2^{(2)}} \cdot \frac{\partial u_2^{(2)}}{\partial z_2^{(1)}} \\ &= 0.741 \cdot 0.187 \cdot 0.45 + -0.218 \cdot 0.176 \cdot 0.55 \\ &= 0.041\end{aligned}$$

Por lo que el resultado para $\frac{\partial L}{\partial w_{2,1}}$ queda:

$$\begin{aligned}\frac{\partial L}{\partial w_{2,1}} &= \frac{\partial L}{\partial z_2^{(1)}} \cdot \frac{\partial z_2^{(1)}}{\partial w_2^{(1)}} \cdot \frac{\partial w_2^{(1)}}{\partial w_{2,1}} \\ &= 0.041 \cdot 0.240 \cdot 0.05 \\ &= 0.000496\end{aligned}$$

Finalmente para $\frac{\partial L}{\partial w_{2,2}}$ se repite el mismo procedimiento y se tiene:

$$\frac{\partial L}{\partial w_{2,2}} = \frac{\partial L}{\partial z_2^{(1)}} \cdot \frac{\partial z_2^{(1)}}{\partial w_2^{(1)}} \cdot \frac{\partial w_2^{(1)}}{\partial w_{2,2}}$$

y usando los resultados anteriores:

$$\frac{\partial u_2^{(2)}}{\partial w_{2,2}} = x_2 = 0.1$$

Por lo que el resultado para $\frac{\partial L}{\partial w_{2,1}}$ queda:

$$\begin{aligned}\frac{\partial L}{\partial w_{2,2}} &= \frac{\partial L}{\partial z_2^{(1)}} \cdot \frac{\partial z_2^{(1)}}{\partial w_2^{(1)}} \cdot \frac{\partial w_2^{(1)}}{\partial w_{2,2}} \\ &= 0.041 \cdot 0.240 \cdot 0.1 \\ &= 0.000984\end{aligned}$$

1.3 Actualización de pesos

Para actualizar los pesos, basta simplemente calcular:

$$w_i^* = w_i - \eta \frac{\partial L}{\partial w_i}$$

Utilizando la tasa de aprendizaje $\eta = 0.5$.

$$\begin{aligned}
w_{1,1}^* &= w_{1,1} - 0.5\frac{\partial L}{\partial w_{1,1}} = 0.15 - 0.5 \cdot 0.00043 = 0.1498 \\
w_{1,2}^* &= w_{1,2} - 0.5\frac{\partial L}{\partial w_{1,2}} = 0.2 - 0.5 \cdot 0.00087 = 0.1996 \\
w_{2,1}^* &= w_{2,1} - 0.5\frac{\partial L}{\partial w_{2,1}} = 0.25 - 0.5 \cdot 0.000496 = 0.2498 \\
w_{2,2}^* &= w_{2,2} - 0.5\frac{\partial L}{\partial w_{2,2}} = 0.3 - 0.5 \cdot 0.000984 = 0.2995 \\
u_{1,1}^* &= u_{1,1} - 0.5\frac{\partial L}{\partial u_{1,1}} = 0.4 - 0.5 \cdot 0.08217 = 0.3589 \\
u_{1,2}^* &= u_{1,2} - 0.5\frac{\partial L}{\partial u_{1,2}} = 0.45 - 0.5 \cdot 0.08272 = 0.4086 \\
u_{2,1}^* &= u_{2,1} - 0.5\frac{\partial L}{\partial u_{2,1}} = 0.5 - 0.5 \cdot -0.02275 = 0.5113 \\
u_{2,2}^* &= u_{2,2} - 0.5\frac{\partial L}{\partial u_{2,2}} = 0.55 - 0.5 \cdot -0.0229 = 0.5615
\end{aligned}$$

Finalmente los parámetros actualizados son:

$$\begin{aligned}
w_{1,1}^* &= 0.1498 \\
w_{1,2}^* &= 0.1996 \\
w_{2,1}^* &= 0.2498 \\
w_{2,2}^* &= 0.2995 \\
u_{1,1}^* &= 0.3589 \\
u_{1,2}^* &= 0.4086 \\
u_{2,1}^* &= 0.5113 \\
u_{2,2}^* &= 0.5615
\end{aligned}$$

1.4 Segunda pasada forward propagation

Luego, nuevamente se hace una pasa hacia delante y se computa nuevamente la función de pérdida.

$$\begin{aligned}
Z^{(1)} &= \begin{pmatrix} z_1^{(1)} \\ z_2^{(1)} \end{pmatrix} \\
&= \sigma(W^{*(1)} \cdot X + w_0^{(1)}) \\
&= \sigma\left(\begin{pmatrix} 0.1498 & 0.1996 \\ 0.2498 & 0.2995 \end{pmatrix}\begin{pmatrix} 0.05 \\ 0.1 \end{pmatrix} + \begin{pmatrix} 0.35 \\ 0.35 \end{pmatrix}\right) \\
&= \sigma\left(\begin{pmatrix} 0.0275 \\ 0.0424 \end{pmatrix} + \begin{pmatrix} 0.35 \\ 0.35 \end{pmatrix}\right) \\
&= \sigma\left(\begin{pmatrix} 0.3775 \\ 0.3924 \end{pmatrix}\right) \\
&= \begin{pmatrix} \frac{1}{1+e^{-0.3775}} \\ \frac{1}{1+e^{-0.3924}} \end{pmatrix} \\
&= \begin{pmatrix} 0.593 \\ 0.597 \end{pmatrix}
\end{aligned}$$

El segundo paso hacia delante ocurre desde la capa oculta hasta la de salida, siguiendo la misma lógica anterior se llega a los valores estimados $\hat{Y}$ de la salida.

$$\begin{aligned}
\hat{Y} &= \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \end{pmatrix} \\
&= \sigma(U^* \cdot Z^{(1)} + w_s^{(2)}) \\
&= \sigma\left(\begin{pmatrix} 0.3589 & 0.4086 \\ 0.5113 & 0.5615 \end{pmatrix}\begin{pmatrix} 0.593 \\ 0.597 \end{pmatrix} + \begin{pmatrix} 0.6 \\ 0.6 \end{pmatrix}\right) \\
&= \sigma\left(\begin{pmatrix} 0.4568 \\ 0.6385 \end{pmatrix} + \begin{pmatrix} 0.6 \\ 0.6 \end{pmatrix}\right) \\
&= \sigma\left(\begin{pmatrix} 1.1057 \\ 1.2385 \end{pmatrix}\right) \\
&= \begin{pmatrix} \frac{1}{1+e^{1.1057}} \\ \frac{1}{1+e^{1.2385}} \end{pmatrix} \\
&= \begin{pmatrix} 0.742 \\ 0.775 \end{pmatrix}
\end{aligned}$$

Calculamos nuevamente la pérdida utilizando la fórmula del error cuadrático medio.

$$\begin{aligned}
L &= \frac{1}{2}[(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2] \\
&= \frac{1}{2}[(0.01 - 0.742)^2 + (0.99 - 0.775)^2] \\
&= 0.291
\end{aligned}$$

Se puede notar que disminuye sutilmente el error.