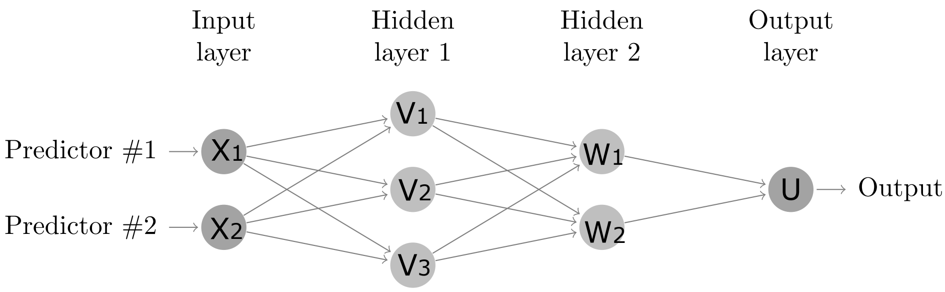
HW 3 MATH 686



$$\begin{cases} a_{11}X_1 & + & a_{12}X_2 & + & d_{11} & = & V_1 \\ a_{21}X_1 & + & a_{22}X_2 & + & d_{12} & = & V_2 \\ a_{31}X_1 & + & a_{32}X_2 & + & d_{13} & = & V_3 \\ b_{11}V_1 + b_{12}V_2 & + & b_{13}V_3 & + & d_{21} & = & W_1 \\ b_{21}V_1 + b_{22}V_2 & + & b_{23}V_3 & + & d_{22} & = & W_2 \\ c_1W_1 & + & c_2W_2 & + & d_3 & = & U \end{cases}$$

1. Find the gradient of U with respect to (W_1, W_2).
2. Find the Jacobian of (W_1, W_2) with respect to (V_1, V_2, V_3)
3. Find the Jacobian of (V_1, V_2, V_3) with respect to $(a_{11}, a_{12}, d_{11}, a_{21}, a_{22}, d_{12}, a_{31}, a_{32}, d_{13})$
4. Treat (X_1, X_2) as constants and find the partial derivatives of U with respect to

$$(a_{11}, a_{12}, a_{21}, a_{22}, a_{31}, a_{32}, d_{11}, d_{12}, d_{13}, b_{11}, b_{12}, b_{13}, b_{21}, b_{22}, b_{23}, d_{21}, d_{22}, c_1, c_2, d_3)$$

First,

$$\frac{\partial U}{\partial d_3} = 1 \quad \begin{pmatrix} \frac{\partial U}{\partial c_1} \\ \frac{\partial U}{\partial c_2} \end{pmatrix} = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}$$

Next, notice that $\frac{\partial U}{\partial b_{11}} = \frac{\partial U}{\partial W_1}\frac{\partial W_1}{\partial b_{11}} + \frac{\partial U}{\partial W_2}\frac{\partial W_2}{\partial b_{11}} = \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right)\begin{pmatrix} \frac{\partial W_1}{\partial b_{11}} \\ \frac{\partial W_2}{\partial b_{11}} \end{pmatrix}$.

In the same way we can derive

$$\frac{\partial U}{\partial b_{12}} = \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right)\begin{pmatrix} \frac{\partial W_1}{\partial b_{12}} \\ \frac{\partial W_2}{\partial b_{12}} \end{pmatrix}$$

$$\frac{\partial U}{\partial b_{13}} = \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right) \begin{pmatrix} \frac{\partial W_1}{\partial b_{13}} \\ \frac{\partial W_2}{\partial b_{13}} \end{pmatrix}$$

$$\frac{\partial U}{\partial b_{21}} = \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right) \begin{pmatrix} \frac{\partial W_1}{\partial b_{21}} \\ \frac{\partial W_2}{\partial b_{21}} \end{pmatrix}$$

$$\frac{\partial U}{\partial b_{22}} = \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right) \begin{pmatrix} \frac{\partial W_1}{\partial b_{22}} \\ \frac{\partial W_2}{\partial b_{22}} \end{pmatrix}$$

$$\frac{\partial U}{\partial b_{23}} = \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right) \begin{pmatrix} \frac{\partial W_1}{\partial b_{23}} \\ \frac{\partial W_2}{\partial b_{23}} \end{pmatrix}$$

$$\frac{\partial U}{\partial d_{21}} = \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right) \begin{pmatrix} \frac{\partial W_1}{\partial d_{21}} \\ \frac{\partial W_2}{\partial d_{21}} \end{pmatrix}$$

$$\frac{\partial U}{\partial d_{22}} = \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right) \begin{pmatrix} \frac{\partial W_1}{\partial d_{22}} \\ \frac{\partial W_2}{\partial d_{22}} \end{pmatrix}$$

Collectively, using matrix multiplication

$$\begin{aligned}
\begin{pmatrix} \frac{\partial U}{\partial b_{11}} \\ \frac{\partial U}{\partial b_{12}} \\ \frac{\partial U}{\partial b_{13}} \\ \frac{\partial U}{\partial b_{21}} \\ \frac{\partial U}{\partial b_{22}} \\ \frac{\partial U}{\partial b_{23}} \end{pmatrix}^{\top} &= \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right) \begin{pmatrix} \frac{\partial W_1}{\partial b_{11}} & \frac{\partial W_1}{\partial b_{12}} & \frac{\partial W_1}{\partial b_{13}} & \frac{\partial W_1}{\partial b_{21}} & \frac{\partial W_1}{\partial b_{22}} & \frac{\partial W_1}{\partial b_{23}} \\ \frac{\partial W_2}{\partial b_{11}} & \frac{\partial W_2}{\partial b_{12}} & \frac{\partial W_2}{\partial b_{13}} & \frac{\partial W_2}{\partial b_{21}} & \frac{\partial W_2}{\partial b_{22}} & \frac{\partial W_2}{\partial b_{23}} \end{pmatrix} \\
&= (c_1, c_2) \begin{pmatrix} V_1 & V_2 & V_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & V_1 & V_2 & V_3 \end{pmatrix} \\
&= \left(c_1 V_1, c_1 V_2, c_1 V_3, c_2 V_1, c_2 V_2, c_2 V_3 \right)
\end{aligned}$$

$$\begin{aligned}
\begin{pmatrix} \frac{\partial U}{\partial d_{21}} \\ \frac{\partial U}{\partial d_{22}} \end{pmatrix}^{\top} &= \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right) \begin{pmatrix} \frac{\partial W_1}{\partial d_{21}} & \frac{\partial W_1}{\partial d_{21}} \\ \frac{\partial W_2}{\partial d_{22}} & \frac{\partial W_2}{\partial d_{22}} \end{pmatrix} \\
&= (c_1, c_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = (c_1, c_2)
\end{aligned}$$

Lastly,

$$
\begin{aligned}
\frac{\partial U}{\partial a_{11}} &= \frac{\partial U}{\partial W_1}\frac{\partial W_1}{\partial a_{11}} + \frac{\partial U}{\partial W_2}\frac{\partial W_2}{\partial a_{11}} \\
&= \frac{\partial U}{\partial W_1}\left(\frac{\partial W_1}{\partial V_1}\frac{\partial V_1}{\partial a_{11}} + \frac{\partial W_1}{\partial V_1}\frac{\partial V_2}{\partial a_{11}} + \frac{\partial W_1}{\partial V_1}\frac{\partial V_3}{\partial a_{11}}\right) \\
&+ \frac{\partial U}{\partial W_2}\left(\frac{\partial W_2}{\partial V_1}\frac{\partial V_1}{\partial a_{11}} + \frac{\partial W_2}{\partial V_1}\frac{\partial V_2}{\partial a_{11}} + \frac{\partial W_2}{\partial V_1}\frac{\partial V_3}{\partial a_{11}}\right) \\
&= \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right)\begin{pmatrix} \frac{\partial W_1}{\partial V_1} & \frac{\partial W_1}{\partial V_2} & \frac{\partial W_1}{\partial V_3} \\ \frac{\partial W_2}{\partial V_2} & \frac{\partial W_2}{\partial V_2} & \frac{\partial W_2}{\partial V_3} \end{pmatrix}\begin{pmatrix} \frac{\partial V_1}{\partial a_{11}} \\ \frac{\partial V_2}{\partial a_{11}} \\ \frac{\partial V_3}{\partial a_{11}} \end{pmatrix} \\
&= (c_1, c_2)\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}\begin{pmatrix} X_1 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

and in general

$$
\begin{aligned}
\frac{\partial U}{\partial a_{ij}} &= \frac{\partial U}{\partial W_1}\frac{\partial W_1}{\partial a_{ij}} + \frac{\partial U}{\partial W_2}\frac{\partial W_2}{\partial a_{ij}} \\
&= \frac{\partial U}{\partial W_1}\left(\frac{\partial W_1}{\partial V_1}\frac{\partial V_1}{\partial a_{ij}} + \frac{\partial W_1}{\partial V_1}\frac{\partial V_2}{\partial a_{ij}} + \frac{\partial W_1}{\partial V_1}\frac{\partial V_3}{\partial a_{ij}}\right) \\
&+ \frac{\partial U}{\partial W_2}\left(\frac{\partial W_2}{\partial V_1}\frac{\partial V_1}{\partial a_{ij}} + \frac{\partial W_2}{\partial V_1}\frac{\partial V_2}{\partial a_{ij}} + \frac{\partial W_2}{\partial V_1}\frac{\partial V_3}{\partial a_{ij}}\right) \\
&= \left(\frac{\partial U}{\partial W_1}, \frac{\partial U}{\partial W_2}\right)\begin{pmatrix} \frac{\partial W_1}{\partial V_1} & \frac{\partial W_1}{\partial V_2} & \frac{\partial W_1}{\partial V_3} \\ \frac{\partial W_2}{\partial V_2} & \frac{\partial W_2}{\partial V_2} & \frac{\partial W_2}{\partial V_3} \end{pmatrix}\begin{pmatrix} \frac{\partial V_1}{\partial a_{ij}} \\ \frac{\partial V_2}{\partial a_{ij}} \\ \frac{\partial V_3}{\partial a_{ij}} \end{pmatrix},
\end{aligned}
$$

Then

$$
\begin{aligned}
\begin{pmatrix} \frac{\partial U}{\partial a_{11}} \\ \frac{\partial U}{\partial a_{12}} \\ \frac{\partial U}{\partial a_{13}} \\ \frac{\partial U}{\partial a_{21}} \\ \frac{\partial U}{\partial a_{22}} \\ \frac{\partial U}{\partial a_{23}} \end{pmatrix}^{\top} &= (c_1, c_2)\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}\begin{pmatrix} X_1 & X_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & X_1 & X_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & X_1 & X_2 \end{pmatrix} \\
&= ((c_1 b_{11} + c_2 b_{21})X_1, \quad (c_1 b_{11} + c_2 b_{21})X_2, \quad (c_1 b_{12} + c_2 b_{22})X_1, \\
&\qquad (c_1 b_{12} + c_2 b_{22})X_2, \quad (c_1 b_{13} + c_2 b_{23})X_1, \quad (c_1 b_{13} + c_2 b_{23})X_2)
\end{aligned}
$$

and similarly

$$
\begin{pmatrix} \frac{\partial U}{\partial d_{11}} \\ \frac{\partial U}{\partial d_{12}} \\ \frac{\partial U}{\partial d_{13}} \end{pmatrix}^{\top} = (c_1, c_2) \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
$$

$$
= (c_1 b_{11} + c_2 b_{21}, c_1 b_{12} + c_2 b_{22}, c_1 b_{13} + c_2 b_{23})
$$

5. Implement the neural networks by finishing the backward propagation function of the R code in "a simple linear neural network (need finish the backward propagation function".

   - Run the entire code and report the testing MSE.
   - Submit your code for the backward propagation function in print.