# CS 1678/2078 HW Backprop


**Abstract**

In this assignment you will be computing the gradients of the weights of multi-layered neural network by hand. This serves as a precursor to the next homework where you will be implementing backprop to train a multi-layered neural network. To submit this assignment, upload a `.pdf` to Gradescope containing your responses to the questions below. You are required to use LaTeX for your write up.


## 1 Partial Derivatives With a Single Sample (34 points)

Consider a neural network with two hidden layers and a linear output layer. The input to the network is a vector of length four, the first hidden layer has three hidden units, the second layer has two, and the last layer as a single unit. Each hidden layer uses the ReLU activation function.

For a single input $x$ and target value $y \in \mathbb{R}$, the loss function for the network is

$$l(\theta) = \frac{1}{2}\left(f(x,\theta) - y\right)^2,$$

with each layer computing,

$$h^i = f^i(h^{i-1}, W^i) = \sigma\left(h^{i-1}{W^i}^\top\right),$$

where $h^i \in \mathbb{R}^{1\times n_i}$ and $W^i \in \mathbb{R}^{n_i \times n_{i-1}}$. Note that we dropped the dataset $D$ in notation for the loss function $l_D(\theta)$. This just makes notation simpler for the assignment. Let the partial derivative of the loss with respect to $f(x,\theta)$ be $\delta$, e.g.

$$\delta = \frac{\partial l(\theta)}{\partial f(x,\theta)} = f(x,\theta) - y$$

1. What is the partial derivative of $l(\theta)$ with respect to the weight $W^3_{1,1}$?

$$\begin{aligned}\frac{\partial l(\theta)}{\partial W^3_{1,1}} &= \frac{\partial h^3}{\partial W^3_{1,1}}\frac{\partial l(\theta)}{\partial h^3} \\ &= \end{aligned}$$

2. What is the partial derivative of $l(\theta)$ with respect to the weight $W^3_{1,2}$?

$$\frac{\partial l(\theta)}{\partial W^3_{1,2}} =$$

3. What are the partial derivatives of $l(\theta)$ with respect to $W^3$.

$$\begin{aligned}\frac{\partial l(\theta)}{\partial W^3} &= \begin{bmatrix} \frac{\partial l(\theta)}{\partial W^3_{1,1}} & \frac{\partial l(\theta)}{\partial W^3_{1,2}} \end{bmatrix} \\ &= \end{aligned}$$

4. What are the partial derivatives of $l(\theta)$ with respect to $h^2_{1,1}$.

$$\frac{\partial l(\theta)}{\partial h^2_{1,1}} =$$

5. What are the partial derivatives of $l(\theta)$ with respect to $h^2$.

$$\frac{\partial l(\theta)}{\partial h^2} = \begin{bmatrix} \frac{\partial l(\theta)}{\partial h^2_{1,1}} & \frac{\partial l(\theta)}{\partial h^2_{1,2}} \end{bmatrix}$$
$$= \delta \begin{bmatrix} ? & ? \end{bmatrix}$$

6. What is the derivative for the ReLU activation function $\sigma(x) = \max(x, 0)$? You can use the notation that $x > y$ evaluates to 1 if true and 0 if false.

$$\frac{d\sigma(x)}{dx} = ?$$

7. What are the partial derivatives with respect to $W^2_{1,j}$ for $h^2_{1,1} = f^2_1(h^1, W^2)$? You may use $z^i = h^{i-1}{W^i}^\top$ and $z^i_{1,1} = h^{i-1}{W^i_{1,\cdot}}^\top$ to simplify your answer.

$$\frac{\partial h^2_{1,1}}{\partial W^2_{1,j}} =$$

8. What are the partial derivatives with respect to $W^2_{2,j}$ for $h^2_{1,1} = f^2_1(h^1, W^2)$?

$$\frac{\partial h^2_{1,1}}{\partial W^2_{2,j}} =$$

9. What are the partial derivatives with respect to $W^2$ for $h^2_{1,1} = f^2_1(h^1, W^2)$?

$$\frac{\partial h^2_{1,1}}{\partial W^2} = \begin{bmatrix} \frac{\partial h^2_{1,1}}{\partial W^2_{1,1}} & \frac{\partial h^2_{1,1}}{\partial W^2_{1,2}} & \frac{\partial h^2_{1,1}}{\partial W^2_{1,3}} \\ \frac{\partial h^2_{1,1}}{\partial W^2_{2,1}} & \frac{\partial h^2_{1,1}}{\partial W^2_{2,2}} & \frac{\partial h^2_{1,1}}{\partial W^2_{2,3}} \end{bmatrix}$$

10. What are the partial derivatives of $l(\theta)$ with respect to $W^2_{i,j}$? Note that using scalar notation we express $h^3$ as

$$h^3_{1,1} = \sum_{q=1}^{n_2} h^2_{1,q} W^3_{1,q} = \sum_{q=1}^{n_2} \sigma\left(\sum_{r=1}^{n_1} h^1_{1,r} W^2_{q,r}\right) W^3_{1,q}.$$

You can use this expression as a starting point for the derivative if you are not comfortable with linear algebra.

$$\frac{\partial l(\theta)}{\partial W^2_{i,j}} =$$

11. What are the partial derivatives with respect to $W^2$ for $l(\theta)$?

$$\frac{\partial l(\theta)}{\partial W^2} = \begin{bmatrix} \frac{\partial l(\theta)}{\partial W^2_{1,1}} & \frac{\partial l(\theta)}{\partial W^2_{1,2}} & \frac{\partial l(\theta)}{\partial W^2_{1,3}} \\ \frac{\partial l(\theta)}{\partial W^2_{2,1}} & \frac{\partial l(\theta)}{\partial W^2_{2,2}} & \frac{\partial l(\theta)}{\partial W^2_{2,3}} \end{bmatrix} = \begin{bmatrix} ? & ? & ? \\ ? & ? & ? \end{bmatrix}$$

12. What are the partial derivatives of $h^2_{1,1}$ with respect to $h^1_{1,j}$?

$$\frac{\partial h^2_1}{\partial h^1_{1,j}} =$$

13. What are the partial derivatives of $h^2_{1,i}$ with respect to $h^1$?

$$\frac{\partial h^2_{1,i}}{\partial h^1} = \begin{bmatrix} ? & ? & ? \end{bmatrix}$$

14. What are the partial derivatives of $l(\theta)$ with respect to $h^1_{1,j}$?

$$\frac{\partial l(\theta)}{\partial h^1_{1,j}} =$$

15. What are the partial derivatives of $l(\theta)$ with respect to $h^1$?

$$\begin{aligned}\frac{\partial l(\theta)}{\partial h^1} &= \begin{bmatrix} \frac{\partial l(\theta)}{\partial h^1_{1,1}} & \frac{\partial l(\theta)}{\partial h^1_{1,2}} & \frac{\partial l(\theta)}{\partial h^1_{1,3}} \end{bmatrix} \\ &= \begin{bmatrix} ? & ? & ? \end{bmatrix}\end{aligned}$$

16. What is the partial derivative of $l(\theta)$ with respect to $W^1_{i,j}$?

$$\frac{\partial l(\theta)}{\partial W^1_{i,j}} =$$

17. What are the partial derivatives of $l(\theta)$ with respect to $W^1$? For conciseness you may use leave your answer in terms of $\frac{\partial l(\theta)}{\partial h^2_{1,j}}$. For further ease of notation, you can write these partial derivatives as $\partial_{h^2_{1,j}} l(\theta)$.

$$\begin{aligned}\frac{\partial l(\theta)}{\partial W^1} &= \begin{bmatrix} \frac{\partial l(\theta)}{\partial W^1_{1,1}} & \frac{\partial l(\theta)}{\partial W^1_{1,2}} & \frac{\partial l(\theta)}{\partial W^1_{1,3}} & \frac{\partial l(\theta)}{\partial W^1_{1,4}} \\ \frac{\partial l(\theta)}{\partial W^1_{2,1}} & \frac{\partial l(\theta)}{\partial W^1_{2,2}} & \frac{\partial l(\theta)}{\partial W^1_{2,3}} & \frac{\partial l(\theta)}{\partial W^1_{2,4}} \\ \frac{\partial l(\theta)}{\partial W^1_{3,1}} & \frac{\partial l(\theta)}{\partial W^1_{3,2}} & \frac{\partial l(\theta)}{\partial W^1_{3,3}} & \frac{\partial l(\theta)}{\partial W^1_{3,4}} \end{bmatrix} \\ &= \begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \end{bmatrix}\end{aligned}$$

## 2 Partial Derivatives for a Batch of Data (16 points)

Instead of computing derivatives for a single data point at a time, it is faster to compute a derivatives for a mini-batch of $m$ data points. First consider a mini-batch size of $m = 2$, e.g., $x \in \mathbb{R}^{2\times 4}$, $y \in \mathbb{R}^{2\times 1}$, $h^1 \in \mathbb{R}^{2\times 3}$, $h^2 \in \mathbb{R}^{2\times 2}$, $h^3 \in \mathbb{R}^{2\times 1}$. Let

$$l_k(\theta) = \frac{1}{2}\left(h^3_{k,1} - y_{k,1}\right)^2.$$

The loss function is now

$$l(\theta) = \frac{1}{m}\sum_{k=1}^{m} l_k(\theta) = \frac{1}{2}\frac{1}{m}\sum_{k=1}^{m}\left(h^3_{k,1} - y_{k,1}\right)^2.$$

1. What is the partial derivative of $l(\theta)$ with respect to $h^3 = f(x, \theta)$? Express your final answer using vector notation.

$$\delta = \frac{\partial l(\theta)}{\partial h^3} = \begin{bmatrix} \frac{\partial l(\theta)}{\partial h^3_{1,1}} \\ \frac{\partial l(\theta)}{\partial h^3_{2,1}} \end{bmatrix}$$
$$=$$

2. What is the partial derivative of $l(\theta)$ with respect to $W^3_{1,1}$?

$$\frac{\partial l(\theta)}{\partial W^3_{1,1}} = ?$$

3. What are the partial derivatives of $l(\theta)$ with respect to $W^3$? Express the final answer using vector notation.

$$\frac{\partial l(\theta)}{\partial W^3} = \begin{bmatrix} \frac{\partial l(\theta)}{\partial W^3_{1,1}} & \frac{\partial l(\theta)}{\partial W^3_{1,2}} \end{bmatrix}$$
$$=$$

4. What are the partial derivatives $l(\theta)$ with respect to $h^2_{\cdot,1}$? Express the final answer using vector notation.

$$\frac{\partial l(\theta)}{\partial h^2_{\cdot,1}} = \begin{bmatrix} \frac{\partial l(\theta)}{\partial h^2_{1,1}} \\ \frac{\partial l(\theta)}{\partial h^2_{2,1}} \end{bmatrix}$$
$$=$$

5. What are the partial derivatives $l(\theta)$ with respect to $h^2$? Express the final answer using vector notation.

$$\frac{\partial l(\theta)}{\partial h^2} = \begin{bmatrix} \frac{\partial l(\theta)}{\partial h^2_{1,1}} & \frac{\partial l(\theta)}{\partial h^2_{1,2}} \\ \frac{\partial l(\theta)}{\partial h^2_{2,1}} & \frac{\partial l(\theta)}{\partial h^2_{2,2}} \end{bmatrix}$$
$$=$$

6. What is the partial derivative of $l(\theta)$ with respect to $W^2_{i,j}$?

$$\frac{\partial l(\theta)}{\partial W^2_{i,j}} =$$

7. What are the partial derivatives of $l(\theta)$ with respect to $W^2$? Express your answer using vector notation.

$$\frac{\partial l(\theta)}{\partial W^2} = \begin{bmatrix} \frac{\partial l(\theta)}{\partial W^2_{1,1}} & \frac{\partial l(\theta)}{\partial W^2_{1,2}} & \frac{\partial l(\theta)}{\partial W^2_{1,3}} \\ \frac{\partial l(\theta)}{\partial W^2_{2,1}} & \frac{\partial l(\theta)}{\partial W^2_{2,2}} & \frac{\partial l(\theta)}{\partial W^2_{2,3}} \end{bmatrix}$$
$$=$$

8. What are the partial derivatives of $l(\theta)$ with respect to $h^1$? Express your answer using vector notation. You can use $\partial_{h^2} l(\theta) = \frac{\partial l(\theta)}{\partial h^2}$ to simplify your answer.

$$\begin{aligned} \frac{\partial l(\theta)}{\partial h^1} &= \begin{bmatrix} \frac{\partial l(\theta)}{\partial h^1_{1,1}} & \frac{\partial l(\theta)}{\partial h^1_{1,2}} & \frac{\partial l(\theta)}{\partial h^1_{1,3}} \\ \frac{\partial l(\theta)}{\partial h^1_{2,1}} & \frac{\partial l(\theta)}{\partial h^1_{2,2}} & \frac{\partial l(\theta)}{\partial h^1_{2,3}} \end{bmatrix} \\ &= \end{aligned}$$