A Brief Introduction to Deep Neural Networks

Sarper Yurtseven
sarperyn@gmail.com

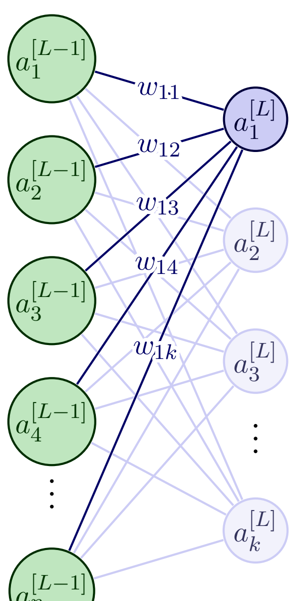
Introduction

One should have a basic knowledge of linear algebra and calculus to understand some concepts in neural networks. If you don't have any prior knowledge, you can have a basic understanding by looking at the notes on systems of linear equations and matrix operations. These notes are not complete yet. However, they are sufficient for the concepts below. (Calculus is not included.)

Forward Propagation

Assume that we have **m** examples. Let's review the general formulas we know.

$n^{[L]}$: number of units in layer L
$A^{[L]} = \sigma(Z^{[L]}) = \sigma(W^{[L]} \cdot A^{[L-1]} + b^{[L]})$
$\mathrm{W}^{[L]}.shape = (n^{[L]}, n^{[L-1]})$
$\mathrm{b}^{[L]}.shape = (n^{[L]}, 1)$
$\mathrm{A}^{[L]}.shape = Z^{[L]}.shape = (n^{[L]}, n^{[L-1]})$



$$\begin{aligned}
a_1^{[L]} &= \sigma\left(z_1^{[L]}\right) \\
&= \sigma\left(w_{11} \cdot a_1^{[L-1]} + w_{12} \cdot a_2^{[L-1]} + \ldots + w_{1n} \cdot a_n^{[L-1]} + b_1^{[L]}\right) \\
&= \sigma\left(\sum_{i=1}^{n} w_{1i} \cdot a_i^{[L-1]} + b_1^{[L-1]}\right)
\end{aligned}$$

$$\begin{aligned}
a_2^{[L]} &= \sigma\left(z_2^{[L]}\right) \\
&= \sigma\left(w_{21} \cdot a_1^{[L-1]} + w_{22} \cdot a_2^{[L-1]} + \ldots + w_{2n} \cdot a_n^{[L-1]} + b_2^{[L]}\right) \\
&= \sigma\left(\sum_{i=1}^{n} w_{2i} \cdot a_i^{[L-1]} + b_2^{[L-1]}\right)
\end{aligned}$$

$$\begin{aligned}
a_3^{[L]} &= \sigma\left(z_2^{[L]}\right) \\
&= \sigma\left(w_{31} \cdot a_1^{[L-1]} + w_{32} \cdot a_2^{[L-1]} + \ldots + w_{3n} \cdot a_n^{[L-1]} + b_3^{[L]}\right) \\
&= \sigma\left(\sum_{i=1}^{n} w_{3i} \cdot a_i^{[L-1]} + b_3^{[L-1]}\right)
\end{aligned}$$

$$\vdots$$

$$\begin{aligned}
a_m^{[L]} &= \sigma\left(z_m^{[L]}\right) \\
&= \sigma\left(w_{m1} \cdot a_1^{[L-1]} + w_{m2} \cdot a_2^{[L-1]} + \ldots + w_{mn} \cdot a_n^{[L-1]} + b_m^{[L]}\right) \\
&= \sigma\left(\sum_{i=1}^{n} w_{3i} \cdot a_i^{[L-1]} + b_k^{[L-1]}\right)
\end{aligned}$$

Getting Matrix Dimensions

From above operations we know:

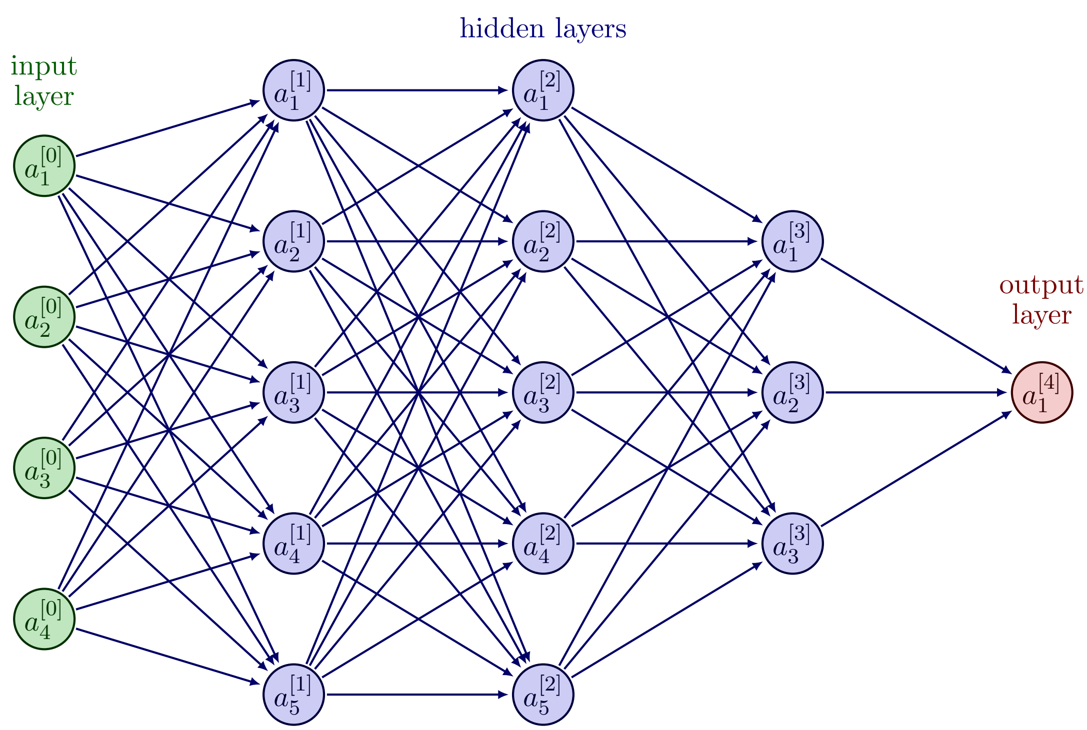
$$
\begin{gathered}
w_{11} \cdot a_1^{[L-1]} + w_{12} \cdot a_2^{[L-1]} + \ldots + w_{1n} \cdot a_n^{[L-1]} + b_1^{[L-1]} = z_1^{[L]} \\
w_{21} \cdot a_1^{[L-1]} + w_{22} \cdot a_2^{[L-1]} + \ldots + w_{2n} \cdot a_n^{[L-1]} + b_1^{[L-1]} = z_2^{[L]} \\
w_{31} \cdot a_1^{[L-1]} + w_{32} \cdot a_2^{[L-1]} + \ldots + w_{3n} \cdot a_n^{[L-1]} + b_1^{[L-1]} = z_3^{[L]} \\
\vdots \\
w_{m1} \cdot a_1^{[L-1]} + w_{m2} \cdot a_2^{[L-1]} + \ldots + w_{mn} \cdot a_n^{[L-1]} + b_1^{[L-1]} = z_k^{[L]}
\end{gathered}
$$

The equations look like a system of linear equations from linear algebra. To get the dimensions, we're gonna treat them like it. Let's rewrite the system as matrices.

$$
\begin{bmatrix}
w_{1,1} & w_{1,2} & \dots & w_{1,n} \\
w_{2,1} & w_{2,2} & \dots & w_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
w_{k,1} & w_{k,2} & \dots & w_{k,n}
\end{bmatrix}
\begin{bmatrix}
a_1^{[L-1]} \\ a_2^{[L-1]} \\ \vdots \\ a_n^{[L-1]}
\end{bmatrix}
+
\begin{bmatrix}
b_1^{[L]} \\ b_2^{[L]} \\ \vdots \\ b_k^{[L]}
\end{bmatrix}
=
\begin{bmatrix}
z_1^{[L]} \\ z_2^{[L]} \\ \vdots \\ z_k^{[L]}
\end{bmatrix}
$$

$$
\underset{k\times n}{W^{[L]}} \times \underset{n\times m}{A^{[L-1]}} + \underset{k\times 1}{b^{[L]}} = \underset{k\times m}{\mathrm{Z}^{[\mathrm{L}]}}
$$

Backward Propagation

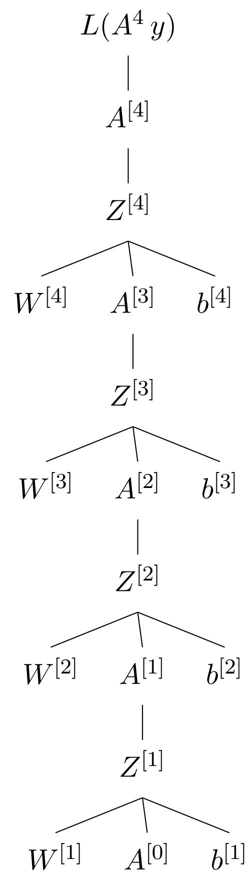


We know that in the first step of backward prop, we're given $dA^{[L]}$. Using $dA^{[L]}$ and the given formulas below, we calculate $dA^{[L-1]}, dW^{[L]}$ and $db^{[L]}$. However, now we gonna just do two of them like we did in the meeting. I'm planning to update this document and add other calculations as well.

Formulas

$$dZ^{[L]} = dA^{[L]} \cdot \sigma(Z^{[L]})$$
$$dW^{[L]} = \frac{1}{m} dZ^{[L]} \cdot A^{[L-1]T}$$

Remember the tree structure.



Instead of Andrew's notation, we're gonna use the following notation. $\frac{\partial L}{\partial Z^{[L]}} = dZ^{[L]}$, $\frac{\partial L}{\partial W^{[L]}} = dW^{[L]}$. Now let's take some derivative

$$\underbrace{\frac{\partial L}{\partial Z^{[4]}}} = \underbrace{\frac{\partial L}{\partial A^{[4]}}} \cdot \underbrace{\frac{\partial A^{[4]}}{\partial Z^{[4]}}}$$

$$\underbrace{\frac{\partial L}{\partial Z^{[3]}}} = \underbrace{\frac{\partial L}{\partial A^{[4]}} \cdot \frac{\partial A^{[4]}}{\partial Z^{[4]}} \cdot \frac{\partial Z^{[4]}}{\partial A^{[3]}}} \cdot \underbrace{\frac{\partial A^{[3]}}{\partial Z^{[3]}}}$$

$$\underbrace{\frac{\partial L}{\partial Z^{[2]}}} = \underbrace{\frac{\partial L}{\partial A^{[4]}} \cdot \frac{\partial A^{[4]}}{\partial Z^{[4]}} \cdot \frac{\partial Z^{[4]}}{\partial A^{[3]}} \cdot \frac{\partial A^{[3]}}{\partial Z^{[3]}} \cdot \frac{\partial Z^{[3]}}{\partial A^{[2]}}} \underbrace{\frac{\partial A^{[2]}}{\partial Z^{[2]}}}$$

$$\underbrace{\frac{\partial L}{\partial W^{[4]}}} = \underbrace{\frac{\partial L}{\partial A^{[4]}} \cdot \frac{\partial A^{[4]}}{\partial Z^{[4]}}} \cdot \underbrace{\frac{\partial Z^{[4]}}{\partial W^{[4]}}}$$

$$\underbrace{\frac{\partial L}{\partial W^{[3]}}} = \underbrace{\frac{\partial L}{\partial A^{[4]}} \cdot \frac{\partial A^{[4]}}{\partial Z^{[4]}} \cdot \frac{\partial Z^{[4]}}{\partial A^{[3]}} \cdot \frac{\partial A^{[3]}}{\partial Z^{[3]}}} \cdot \underbrace{\frac{\partial Z^{[3]}}{\partial W^{[3]}}}$$

$$\underbrace{\frac{\partial L}{\partial W^{[2]}}} = \underbrace{\frac{\partial L}{\partial A^{[4]}} \cdot \frac{\partial A^{[4]}}{\partial Z^{[4]}} \cdot \frac{\partial Z^{[4]}}{\partial A^{[3]}} \cdot \frac{\partial A^{[3]}}{\partial Z^{[3]}} \cdot \frac{\partial Z^{[3]}}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial Z^{[2]}}} \cdot \underbrace{\frac{\partial Z^{[2]}}{\partial W^{[2]}}}$$