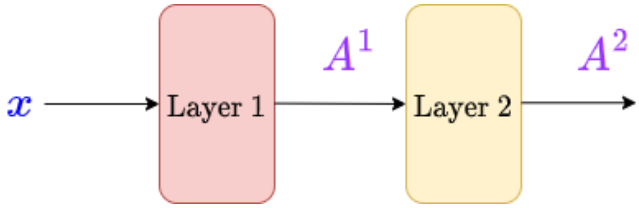
Explanatory Notes for 6.390

Shaunticlair Ruiz (Current TA)

Fall 2022

Adding a second layer

So, let's add one more **layer**. We'll label layers by using a **superscript**: W^1 is the set of **weights** for the **first** layer, for example.



We have two separate outputs: A^1 and A^2.

Clarification 1

Superscripts in our notation indicate the **layer** that our value is associated with.

They do **not** represent exponentiation!

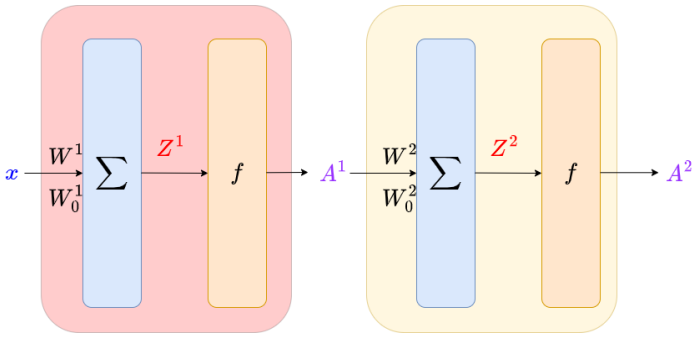
Example: Z^3 would be the **pre-activation** for layer 3: it is **not** Z "cubed".

What can we learn from this?

- The **output** of layer 1, A^1, is the **input** to layer 2.
- Thus, the output dimension n^1 of layer 1 must **match** the input m^2 of layer 2:

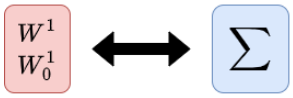
$$n^1 = m^2 \tag{1}$$

Let's break these into their components again.



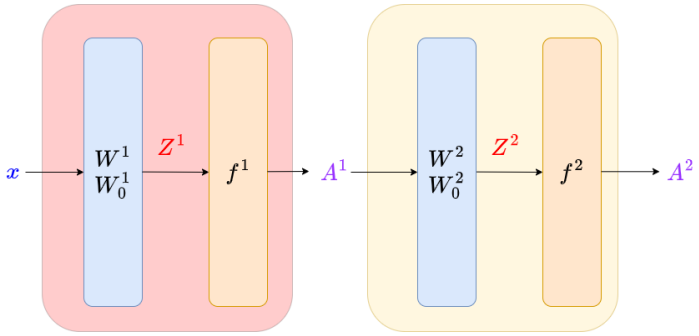
We have two separate outputs: A^1 and A^2.

To distinguish between the linear functions in each layer, we'll just notate them using the weights and offsets.



These two are equivalent (if in the same layer)! We'll use the notation on the left, so that you know which layer our unit is in.

And this gives us:



Now, we can make our functions. For layer one:

$$A^1 = f(Z^1) = f\left((W^1)^T x + W_0^1\right) \tag{2}$$

And layer two:

$$A^2 = f(Z^2) = f\left((W^2)^T A^1 + W_0^2\right) \tag{3}$$

We can use this to build our **general** pattern.