# Explanatory Notes for 6.390

Shaunticlair Ruiz (Current TA)

Fall 2022

## Example: Linear Regression

Let's go through some **examples**. We mentioned in the **beginning** of this chapter that our neuron could be most of the simple **models** we've worked with.

So, let's give that a go: we'll start by doing **linear regression**.

$$h(x) = \theta^T x + \theta_0$$

This model is exclusively **linear**: we just have to replace $\theta$ with $w$.

$$z(x) = w^T x + w_0$$

So, our linear component is **done**: $(\theta, \theta_0) = (w, w_0)$.

What about our **activation** function?

Well, activation allows for **nonlinear** functions. But, we don't **want** to make it nonlinear.

In fact, we've already got what we **want**: we don't want the **activation** to do anything at **all**.

So, we'll use **this** function:

**Concept 1**

The **identity function** $f(z)$ is a function that has no **effect** on your **input**.

$$f(z) = z$$

By "having no effect", we mean that the input is **unchanged**: this is true even if your input is **another function**:

$$f(g(x)) = g(x) \tag{1}$$

We call it the "identity" because the input's identity is unchanged!

So, the **identity** function is our activation function: it keeps our **linearity**.

**Concept 2**

**Linear Regression** can be represented with a **single neuron** where

- We keep our **linear component**, but set $(\theta, \theta_0) = (w, w_0)$.

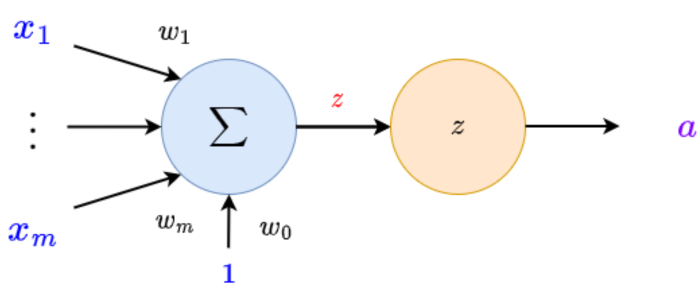
$$z(x) = w^T x + w_0$$

- Our **activation function** is the **identity** function,

$$f(z) = z$$

- Our **loss function** is **quadratic loss**.

$$\mathcal{L}(a, y) = (a - y)^2$$




## Example: Linear Logistic Classifiers

Now, we do the same for LLCs: it's already broken up into **two** parts in our **classification** chapter.

First, the **linear** component. This is the same as linear regression:

$$z = \theta^T x + \theta_0 \tag{2}$$

And then, the **logistic** component:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{3}$$

This second part is **nonlinear**: its our **activation** function!

**Concept 3**

A **Linear Logistic Classifier** can be represented with a **single neuron** where

- We keep our **linear component**, but set $(\theta, \theta_0) = (w, w_0)$.

$$z(x) = w^T x + w_0$$

- Our **activation function** is the **sigmoid** function,

$$f(z) = \sigma(z) = \frac{1}{1 + e^{-z}}$$

- Our **loss function** is **negative-log likelihood** (NLL)

$$\mathcal{L}_{nll}(a, \quad y^{(i)}) = -\left(y^{(i)} \log a + \left(1 - y^{(i)}\right) \log (1 - a)\right)$$