Explanatory Notes for 6.390

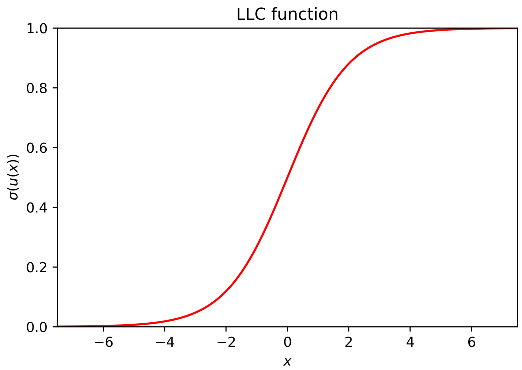
Shaunticlair Ruiz (Current TA)

Fall 2022

Modifying our sigmoid

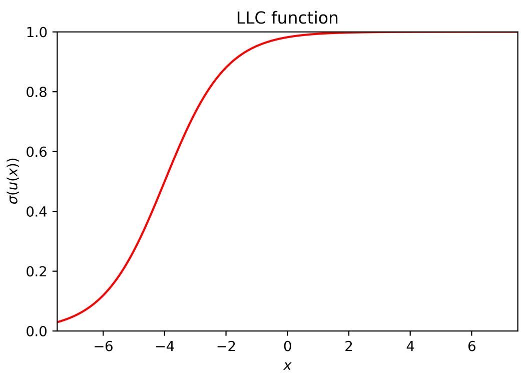
What happens when you modify the **parameters** of an LLC? Let's find out.

We'll use a 1-D input: our variables will be θ (scalar) and θ_0: $\theta x + \theta_0$



Our baseline LLC: $u(x) = 1x + 0$

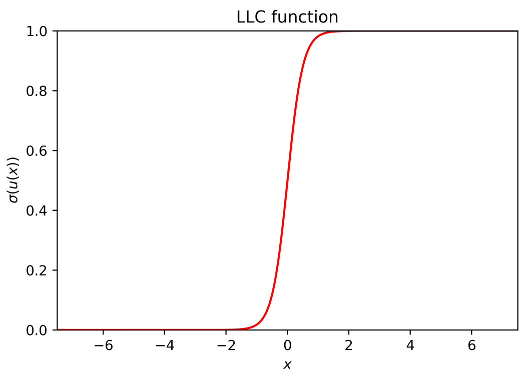
What if we shift by increasing θ_0?



Our shifted LLC: $u(x) = 1x + 4$. θ_0 shifts us along the x-axis!

Just like before, it **shifts** us in the **opposite** direction: if θ_0 is **positive**, we shift in the **negative** direction, and vice versa.

What if we increase the magnitude of θ!



Our new LLC: $u(x) = 4x$. Increasing θ makes our function steeper!

Making the magnitude of θ larger makes our function **change** faster.

This makes some sense: if θ (linear slope of $u(x)$) makes $u(x)$ **change** faster, it will make $\sigma(u)$ change faster **too**.

You can combine these changes as well: you can shift your LLC with θ_0, and also make it steeper/less steep by changing magnitude of θ.

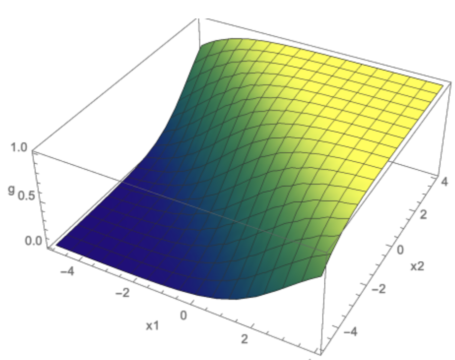
Concept 1

When working with **sigmoids**, you can **transform** them using your **parameters**:

- A higher **magnitude** $\|\theta\|$ makes the slope **steeper**, and answers more **confident**.
- **Increasing** θ_0 **shifts** the sigmoid in the $-\theta$ **direction**, and vice versa.

Viewing our sigmoid in 3D

Let's quickly take a look at a sigmoid in 3D, with two inputs:

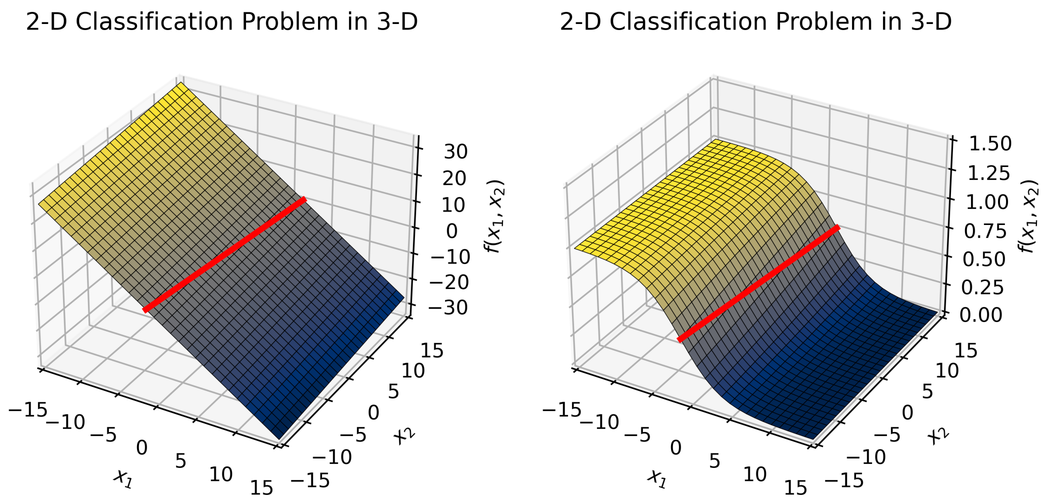


As you can see, you get mostly the same shape: if you look at it from the side, it's exactly the same, in fact! Just stretched out into 3D.

LLCs and LCs have the same boundary

One more important thing to note: noticed that we set $\sigma_{thresh} = .5$, because that was when $u(x) = 0$.

This means that, if our threshold is 0.5, then the boundary of our LLC should look exactly the same as if it were LC: the only difference is the values that we *can't* see:



Despite having different shapes in 3D, they both create 2-D **linear** classifiers: on the left, $u(x) = 0$, and on the right, $\sigma(u) = .5$.

One way to think about this difference is that while one may be logistic, they are both **linear**: they both create the same **linear separator**.

The main benefit of switching to LLC is that $\sigma(u)$ has a useful **gradient**, while $\text{sign}(u)$ does **not**, so we can do **gradient descent**.

The probabilistic interpretation is also more appropriate: we shouldn't be fully confident in our answers.

Even if we adjust our threshold σ_{thresh}, that will simply shift the linear classifier.

Concept 2

LLCs (Linear Logistic Classifiers) and **LCs** (Linear Classifiers) both create a **linear hyperplane separator** in $d - 1$ dimensional **space**.

If the **threshold value** is 0.5, then they have the **exact same** separator.