Explanatory Notes for 6.390

Shaunticlair Ruiz (Current TA)

Fall 2022

IMPORTANT: A difference between regression and classification

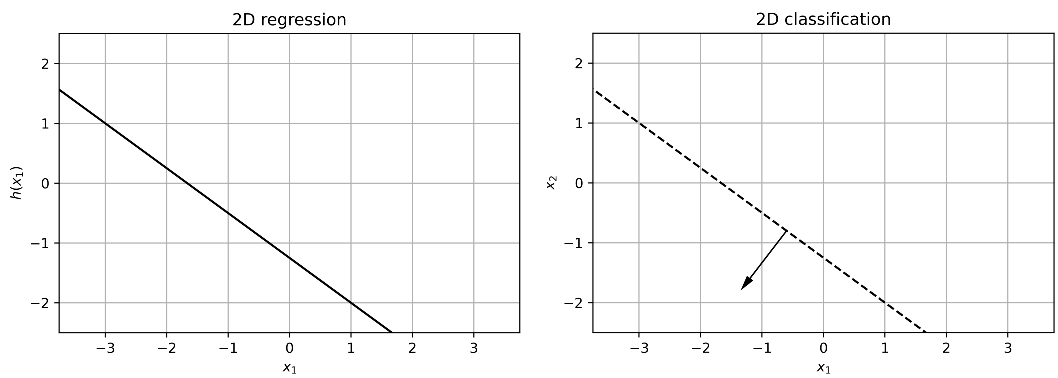
Here is an important misconception that comes up between regression and classification.

Both functions use the equation

$$\theta^T x + \theta_0 \tag{1}$$

So, one might think of them as interchangeable.

However, they are **not**. Why is that?



These two plots look almost the same, but represent completely different things!

Notice that these two plots are **both** plotted in 2-D, and both have a **line** plotted. But, they **aren't** as **similar** as they look.

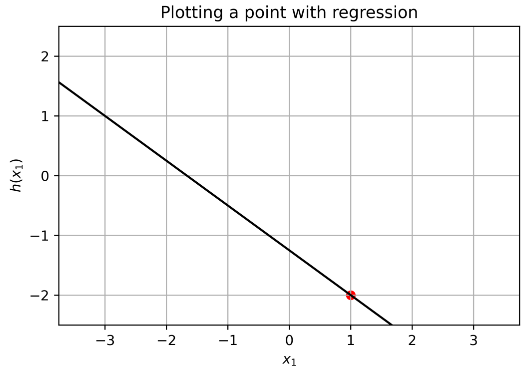
Notice, for example, that the regression plot has **only** x_1, while the classification plot has x_1 **and** x_2.

The reason why? The **output**.

- In **regression**, the output is a **real number**: every point on that line represents an input x_1, and an output $h(x_1)$.
  - This plot can only contain **one** input variable: the **second** axis is reserved for the **output**!
- In **classification**, the output is **binary**. So, that line represents only the **values** where the output is $h(x) = 0$.
  - This plot can contain **two** input variables: x_1 and x_2. Rather than **displaying** the output, we only show one **slice** of the output: the $h(x) = 0$ slice.

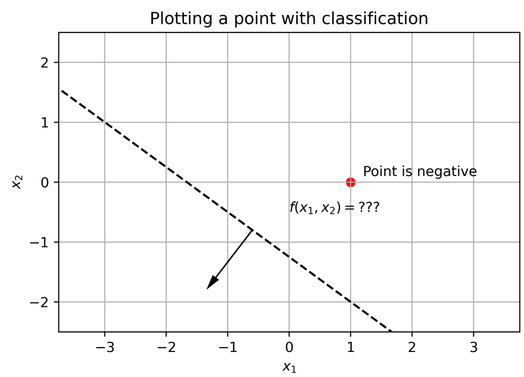
If we think in terms of $f(x) = \theta^T x + \theta_0$, we can compare them directly.

The regression plot shows the exact value on the y-axis. If we want to know what $f(x_1 = 1)$ looks like, we can check the plot: we just get $f(1) = -2$.



We have one input, and we get the exact value of our output.

But the classification plot **doesn't**! We aren't given the value of $\theta^T x + \theta_0$ at $x = (1, 0)$: we just know that it's **negative**.



We have two input, and we **don't** get the exact output.

If we wanted to know the exact value of our 2-D classification, we would need to view it as a plane in 3-D space.

This is the trade-off between these two plots: one gives more information about the output, and the other allows for more inputs in a lower dimension.

Clarification 1

Regression and **classification** plots that look the same, have **different functions**:

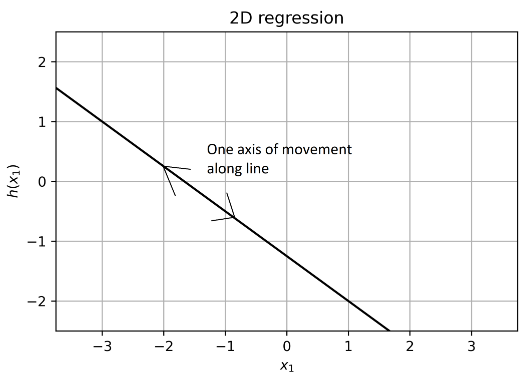
When looking at the output of $f(x) = \theta^T x + \theta_0$,

- A **regression** plot gives the **exact numeric** $f(x)$.
- A **classification** plot only gives the **sign** of the $f(x)$.

When plotting n inputs,

- A **regression** plot uses a $d + 1$ dimensions (d-dim hyperplane) to plot: +1 for the **output**.
- A **classification** plot only needs d dimensions ($(d-1)$-dim hyperplane): we only see the $f(x) = 0$ **hyperplane**.

Why do we need $d + 1$ dimensions to plot a d-dimensional **hyperplane**? You can think of it this way: a **line** in 2-D space is a 1-D **hyperplane**: we have only **one axis** we can move on the line.



Our plot is 2-D, but we can only move along one axis on our line!

Because of these differences, θ also acts differently:

Clarification 2

θ appears differently in 2-D regression and classification:

- In **2-D regression**, θ is the **slope** of the line

$$h(x) = \theta x + \theta_0 \tag{2}$$

- In **2-D classification**, θ is the **normal vector** of the line

$$0 = \theta^T x + \theta_0 \tag{3}$$