Explanatory Notes for 6.390

Shaunticlair Ruiz (Current TA)

Fall 2022

Input Space vs. Parameter Space

One more thing to note: we have two similar situations.

- J is a **function** with θ as an **input**: $J(\theta)$.
- h is a **function** with x as an **input**: $h(x)$.

In both cases, we can imagine the **output** as the "**height**" of our function: the **hill** we mentioned before. This **physical** intuition is useful to **gradient descent**.

But, what about **input** to our function? That's the x-axis our hill is floating above:

- With $h(x)$, our x-axis was our **input space**, all possible x_1 values: the "space" containing all of our possible inputs.
- With $J(\theta)$, our x-axis is the **parameter space**, all possible θ values. We also called this our "**hypothesis space**".

We're assuming 1-D right now for simplicity. If we were 2-D, we'd have an entire 2D grid under our hill!

Definition 1

The **parameter space** is our set of all **possible** parameter combinations.

This is the same as the **hypothesis space**, because our parameters **define** our hypothesis.

When we **optimize** our hypothesis, we are "**exploring**" the hypothesis space.

This also gives us an idea of which hypotheses are "**similar**": those which are **closer** in parameter space (which we used, when we were doing regularization $\|\theta - \theta_{old}\|$).

This is the **space** we're exploring, as we try to move **downhill**.

Clarification 2

Pay attention to your **axes**!

Sometimes, we're doing a 2-D or 3-D plot of J, and our inputs are θ_k. Other times, we're plotting hypothesis h, with our axes x_i.

These two plots could have the same surface, but they **represent** completely different things.