Explanatory Notes for 6.390

Shaunticlair Ruiz (Current TA)

Fall 2022

Including θ_0 in θ

Trying to Simplify

Our approach will involve a lot of **algebra**. Because of that, it's worth it to **simplify** our formula as much as possible beforehand.

$$J(\theta, \theta_0) = \frac{1}{n}\sum_{i=1}^{n}\left(\underbrace{(\theta^T x^{(i)} + \theta_0)}_{\text{guess}} - \underbrace{y^{(i)}}_{\text{answer}}\right)^2 \tag{1}$$

Most parts of this equation can't really be **simplified**: y and x are just variables, and we can't do anything with the **sum** without knowing our data points.

But, one thing that was strange is that we **separated** θ_0 from our other θ_k terms. Maybe we can **fix** that.

Combining θ and θ_0

Let's go back to our **original** equation for $(\theta^T x + \theta_0)$, before we switched to **vectors**.

We drop the $^{(i)}$ notation whenever it isn't necessary, to de-clutter the equations. We only do this when we don't care which data point we're using.

$$h(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots + \theta_d x_d \tag{2}$$

We converted this into a **dot product** because each θ_n term is **multiplied** by an x_k term, except θ_0.

But if we **really** want to include θ_0, then could we? We know what's missing: "θ_0 is **not** multiplied by an x_k term". So... could we get one? Is there a x_0 factor we could **find**?

We need θ_0 to be **multiplied** by something. Is there something we could "**factor** out"? How about: $x_0 = 1$?

You can always factor out 1 without changing the value!

$$h(x) = \theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots + \theta_d x_d \tag{3}$$

So, this means we just have to **append** a 1 to our vector x. At the **same time**, we'll append θ_0 to θ!

$$x = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_d \end{bmatrix}, \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \\ \vdots \\ \theta_d \end{bmatrix}, \qquad h(x) = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \\ \vdots \\ \theta_d \end{bmatrix} \cdot \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_d \end{bmatrix} \tag{4}$$

We'll write that symbolically, and then apply a transpose.

$$h(x) = \theta \cdot x = \theta^T x \tag{5}$$

Concept 1

Sometimes, to simplify our algebra, we can **append** θ_0 to θ.

In order to do this, we have to **append** a value of 1 to x as well.

Once we do this, we can **write**

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

We **have** to append this 1 to every single $x^{(i)}$ in order for this to **work**. But, now we can treat our parameters as **one vector**.