1. Ohm's Law With Noise

We are trying to measure the resistance of a black box. We apply various i_{test} currents and measure the ouput voltage v_{test}. Sometimes, we are quite fortunate to get nice numbers. Oftentimes, our measurement tools are a little bit noisy, and the values we get out of them are not accurate. However, if the noise is completely random, then the effect of it can be averaged out over many samples. Say that we repeat our test many times.

Test	i_{test} (mA)	v_{test} (V)
1	10	21
2	3	7
3	-1	-2
4	5	8
5	-8	-15
6	-5	-11

(a) Plot the measured voltage as a function of the current.

(b) Suppose we stack the currents and voltages to get $\vec{I} = \begin{bmatrix} 10 \\ 3 \\ -1 \\ 5 \\ -8 \\ -5 \end{bmatrix}$ and $\vec{V} = \begin{bmatrix} 21 \\ 7 \\ -2 \\ 8 \\ -15 \\ -11 \end{bmatrix}$. Can you solve for R?

What conditions must $\vec{I}$ and $\vec{V}$ satisfy in order for us to solve for R?

(c) Ideally, we would like to find R such that $\vec{V} = \vec{I}R$. If we cannot do this, we'd like to find a value of R that is the *best* solution possible, in the sense that $\vec{I}R$ is as "close" to $\vec{V}$ as possible. The idea of a best solution is subjective and dependent on the cost function we are using. One way of expressing this cost function in terms of R is to quantify the difference between each component of $\vec{V}$ (V_j) and each component of $\vec{I}R$ (I_jR) and add these "differences" up as follows:

$$\text{cost}(R) = \sum_{j=1}^{6} (V_j - I_jR)^2$$

Do you think this is a good cost function? Why or why not?

(d) Show that you can also express the above cost function in vector form, that is,

$$\text{cost}(R) = \left\langle (\vec{V} - \vec{I}R), (\vec{V} - \vec{I}R) \right\rangle$$

(e) Find $\hat{R}$, the optimal R that minimizes $\text{cost}(R)$.

Hint: Use calculus and minimize the expression in part (c).

(f) On your original IV plot, also plot the line $v = \hat{R}i$. Can you visually see why this line "fits" the data well? What if we had guessed $R = 3$? How well would we have done? What about $R = 1$?

Calculate the cost functions for each of these choices of R to validate your answer.

(g) Now, suppose that we add a new data point: $i_7 = 2\,\text{mA}$, $v_7 = 4\,\text{V}$. Will $\hat{R}$ increase, decrease, or remain the same? Why? What does that say about the line $v = \hat{R}i$?

(h) Let's add another data point: $i_8 = 4\,\text{mA}, v_8 = 11\,\text{V}$. Will $\hat{R}$ increase, decrease, or remain the same? Why? What does that say about the line $v = \hat{R}i$?

(i) Now your mischievous friend has hidden the black box. You want to know what the output voltage would be across the terminals if you applied $5.5\,\text{mA}$ through the black box. What would your best guess be? (This is an example of estimation from machine learning! You have *learned* what is going on inside the black box by making observations, and now you're using what you learned to make estimates.)

2. Demonstration: Triangulation With Noise!

3. Polynomial Fitting

Least squares may seem rather boring at first glance – we're just using it to "solve" systems of linear equations, after all. But, at further glance, it actually comes in a variety of sizes and flavors! For instance, you can solve problems that have decidedly non-linear elements in them using least squares. Let's see how.

Last discussion, we had seen how to "fit" data in the form of (x, y) to a line. This made sense because the input-output relationship was fundamentally linear (Ohm's law).

But what if this relationship was not linear? For instance, the equation of the orbit of a planet around the sun is an ellipse. The equation for the trajectory of a projectile is a parabola. In these sorts of scenarios, how does one fit observation data to the correct curve?

In particular, say we *know* that the output, y, is a *quartic* polynomial in x. This means that we know that y and x are related as follows:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$$

We're also given the following observations:

x	y
0.0	24.0
0.5	6.61
1.0	0.0
1.5	−0.95
2.0	0.07
2.5	0.73
3.0	−0.12
3.5	−0.83
4.0	−0.04
4.5	6.42

(a) What are the unknowns in this question? What are we trying to solve for?

(b) Can you write an equation corresponding to the first observation (x_0, y_0), in terms of a_0, a_1, a_2, a_3, and a_4? What does this equation look like? Is it linear?

(c) Now, write a system of equations in terms of a_0, a_1, a_2, a_3, and a_4 using *all of the observations*.

(d) Finally, solve for a_0, a_1, a_2, a_3, and a_4 using IPython. You have now found the quartic polynomial that best fits the data!

(e) We will now do another example in the IPython notebook and see how to do polynomial fitting quickly using IPython!