Data Fitting

Group 4
Scott Kobos

Overview

So you've made measurements and graphed the results, it's time to talk about fitting. During the theory section, you spend time building a model (ie an equation) that predicts the behavior of the system you are going to measure. For example, for an object in free-fall released from rest, we predicted the object's position should follow

$$y = y_0 - \frac{1}{2}gt^2. \tag{1}$$

An experiment to illustrate this model is to measure the position of a dropped object as a function of time. The end result is a data table of positions (y), uncertainty in position (δy), and time (t). If you make a graph, you can verify that this appears to be an inverted parabola. So far so good, but you want more verification. Let's fit the data.

Define the Fit Function

To define a fit function, you take the equation representing your model and replace all quantities not on the x- or y-axis with constants. For example, Equation 1 would be rewritten as

$$y = c_1 + c_2x^2 \tag{2}$$

where c_1 and $c2$ are constants. When you ask MatLab to fit the data using this model, it will find values for c_1 and c_2 which produce a curve closest to all the data. We expect

$$c_1 = y_0 \tag{3}$$

$$c_2 = -\frac{1}{2}g \tag{4}$$

and can use this information to verify the model.

Fitting the Data

Open the MatLab file "FitExample.m" and run the script. If you would prefer, you can copy the entire text and paste it into a new livescript. The file has comments explaining the purpose of each command. I will give an overview of the process here.

1. First, I define the data that I want to fit. I use general variables x, y, and yerr for the quantity graphed on the x- and y-axis and the array to be used as errorbars.

2. The fit function matches Equation 2 and is stored in the variable *fitf.* The coefficients for the fit function are stored in c. x is the quantity plotted on the x-axis in your fit function.

3. You must make a guess as to what you expect the coefficients will be. The closer the guess, the more likely MatLab is to make a good fit. In the example, I guess for $c_1 = 1.5$ since by inspection I see that the value of y at $t = 0$ is 1.5. I guess that $c_2 = -5$ since $g \approx 10$ m/s^2 and, since $c_2 = -\frac{1}{2}g$, -5 is a reasonable guess.

4. The fit is performed using the function *nlinfit*. You can look up more examples in the online documentation. It takes in everything already defined and produces optimized fit coefficients in the variable *cfit*. The variable *CovB* is a matrix which is used to determine the uncertainty in the fit coefficients.

5. Finally, we calculate the uncertainty in the fit coefficients using the covariant matrix *CovB*.

6. Now, all that is left is to make plots! I make an errorbar plot and then plot the fit function using the optimized coefficients.

After reading through the script, put the coefficients with the errors into the table below and create a figure for the graph.

Coeff	Value
c_1	$1.50 \pm .01m$
c_2	$-4.91 \pm .04\ m/s^2$

Table 1: Coefficients using Equation 2 to fit data in Example 1.

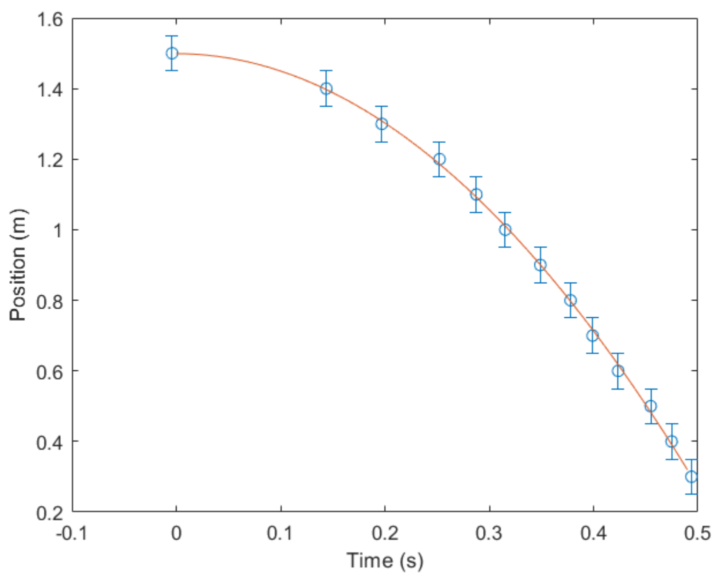

Figure 1: This graph shows the fit between the data of an object in free-fall released from rest and the function of fit.

Calculate Quantities from Coefficients

You now have valuable information in the fit coefficients. Equation 3 tells us that the first coefficient corresponds to the initial object height. Your fit result should well match what you see in the graph: an initial height near 1.5 m. The second fit coefficient is more interesting.

Equation 4 relates to the gravitational constant g. We can use the fit coefficient to calculate a value for g which *should* be near an accepted value of $g = 9.8$ m/s^2. Solving for g we find,

$$g = -2c_2. \tag{5}$$

Use this info to report your experimentally determined values for y_0 and g.

Quantity	Value
y_o	$1.50 \pm .01m$
g	$9.82 \pm .04\ m/s^2$

Table 2: Physical quantities calculated from fit parameters.

Evaluation

Now you have some qualitative and quantitative information to evaluate Equation 1 as a model for the data. Use your graph and your fitted results to answer the questions below.

1. Looking at the line of best fit, how well does it describe the data? Is it near most data points? Is it within errorbars?

   The line of best fit describes the data very well. It is on each data point and within the errorbars.

2. The model predicted a value for y_0 which you have recorded in Table 2. How well does this agree with the data? (Make sure to consider the size of the errorbars) Does it suggest the model is correct or incorrect?

   The predicted value for y_0 is $1.4989 \pm .0059$ this agrees with the data. The data tells us that y_0 is $1.5m$ and the predicted value with the error can be anywhere from 1.493 to 1.5048, which 1.5 is within, suggesting that the model is correct.

3. The model predicted a value for g which you have recorded in Table 2. How well does this agree with the theoretical value? (Make sure to consider the size of the errorbars) Does it suggest the model is correct or incorrect?

   The predicted value for g is within a range from 9.783 to 9.865, this agrees with the theoretical value of 9.8, which falls within that range. This suggests that the model is correct.

Try it Yourself

Now that you've walked through an example, it's time to fit data for yourself. The data in the file "DataSet1.csv" was measured in a nuclear decay experiment. During radioactive decay, the number of unstable nuclei decreases exponentially over time according to the relationship

$$N = N_0 e^{-t/\tau} \tag{6}$$

where N is the number of unstable nuclei at a given time t, N_0 is the initial number of unstable nuclei, and τ is the time constant. Follow the procedure outlined in the previous section to produce a graph of the fit, table of coefficients, and table of quantities showing the fit values for N_0 and τ. Include these in the space below.

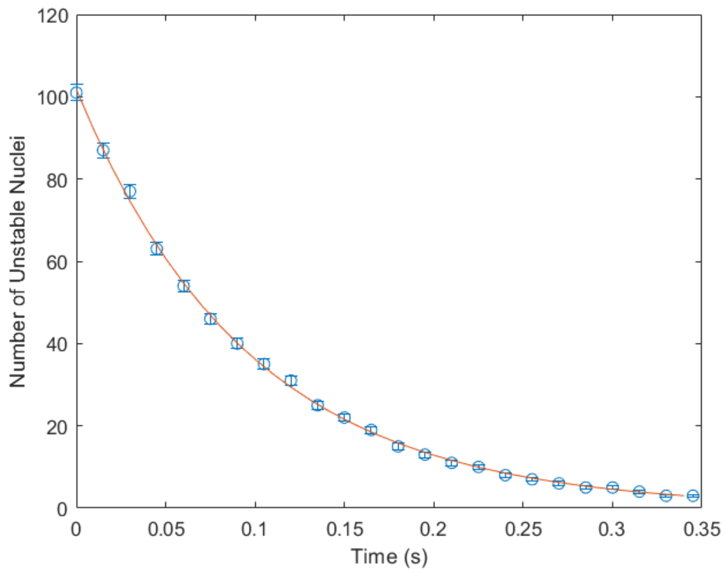

Figure 2: This graph shows the fit between the data of a radioactive decay and the fit function.

Coeff	Value
c_1	$101.7422 \pm .5434$
c_2	$.097 \pm .001s$

Table 3: Coefficients using Equation 6 to fit data.

Good or Bad Fit Function

Now you are going to apply two different fit functions to the same set of data with the goal of determining which function best matches the data. Here's the scenario: In developing the model for a system, you had two different ideas. The first produces a quadratic relationship with the fit function of the form

$$y = c_1 + c_2 x^2. \tag{7}$$

The second is similar but produces a quartic relationship with the fit function of the form

$$y = c_1 + c_2 x^4. \tag{8}$$

You performed an experiment and the data is available in the file "DataSet2.csv" on the moodle page.

Your job is to perform a fit of the data using each fit function, include the graphs of each below and create tables with the coefficients with their error. Then, write a paragraph assessing which of the two models best represents the data measured.

The function that best fits the data from DataSet2 is the quadratic function. The quadratic function passes though each point, or within its error bar, while the quartic function barely passes through a few of the points and their error bars. Also for the quadratic function, the error of the coefficients is smaller for c_1. While for c_2 the error is larger for the quadratic function, both errors are comparably low.

Coeff	Value
N_0	$101.7422 \pm .5434$
τ	$.097 \pm .001s$

Table 4: Fit values for the quantities of $N_0$and τ. Table is for Try it Yourself section, however Overleaf is determined to not allow the logical figure placement within that section.

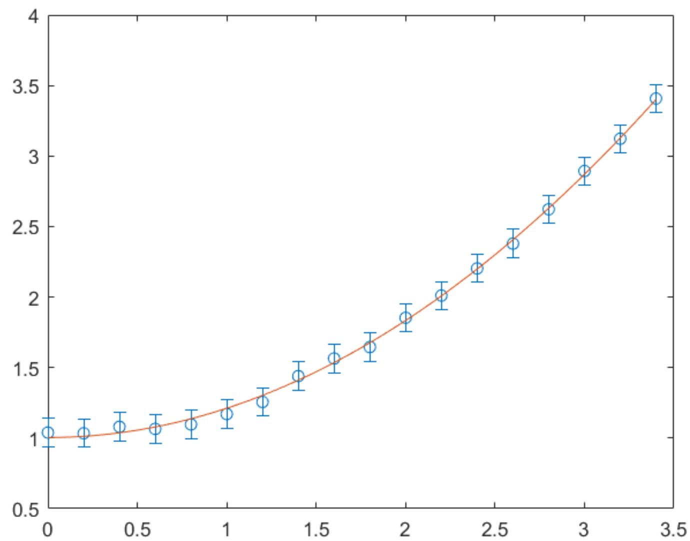

Figure 3: This graph shows the fit of the quadratic function.

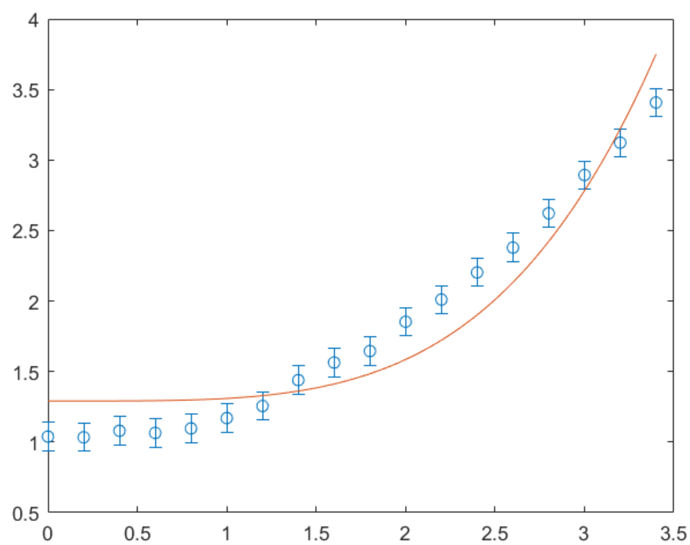

Figure 4: This graph shows the fit of the quartic function.

Coeff	Value
c_1	$1.0 \pm .01m$
c_2	$.2 \pm .001$

Table 5: Coefficients using Equation 7 to fit data.

Quantity	Value
c_1	$1.29 \pm .067m$
c_2	$.02 \pm .001$

Table 6: Coefficients using Equation 8 to fit data.