

Laboratory 2: Falling objects

Hypothesis: The motion of a falling object is characterized by a constant acceleration, a velocity which increases linearly in time, and a position which increases quadratically in time.

Goals: Understanding of linear regression analysis (“trendlines”) and correlation; use of DataStudio software.

1 Introduction

At this point, you should be familiar with the basic equations of one dimensional motion without air resistance. In the present experiment, we will release an object from rest a distance x_o from a motion sensor, and will track its position, velocity, and acceleration. The relevant equations we have developed are, presuming the motion to occur along the x axis,

$$x(t) = x_o + v_o t + \frac{1}{2} a t^2 = x_o \pm \frac{1}{2} g t^2 \quad (1)$$

$$v(t) = v_o + a t = v_o \pm g t \quad (2)$$

$$a = \pm g \quad (3)$$

Here t is time, $x(t)$ is position as a function of time, $v(t)$ is velocity as a function of time, $v_o = 0$ is the initial velocity (the object starts out at rest), a is the acceleration, and $g \approx 9.8 \text{ m/s}^2$ is the acceleration due to gravity. The \pm sign indicates that the acceleration may be $+g$ or $-g$, depending on whether the object falls toward or away from the sensor; which is which is something you will need to figure out during the experiment!

It is the goal of this experiment to verify these relationships and learn a few data analysis techniques that we will use throughout the course.

1.1 Linear regression and trendlines

Let’s say we have a series of data on an $x-y$ plot which seemed to be scattered about a straight line, i.e., the data seem to at least approximately obey a linear relationship. Better yet, perhaps we even have a theoretical model which *predicts* a linear relationship for this particular experimental data. This is the case for the $v(t)$ data you will generate during your experiment.

Linear regression attempts to model the relationship between two variables (say, x and y) by fitting a linear equation (a line) to the observed data. One variable is considered to be an explanatory variable (usually on the x axis), and the other is considered to be a dependent variable (usually on the y axis). The explanatory variable is the thing you are changing or controlling in the experiment, the dependent variable is your primary measurement. For example, in the experiment today you will acquire and plot velocity versus time data which is predicted to obey a linear relationship of the form $v = v_o + a_o t$ where $a_o = -g$ is the constant acceleration and v_o the initial velocity.

If the data *look* linear, and it seems like we could draw a straight line through at least most of the data ... how do we find the best possible line to draw? That is, what is the equation of a straight line that has the minimum total deviation from the actual data set? Further, how well does this line describe our data (i.e., how close is our data to being linear)? The line of best fit is often called a “fit” or “trendline,” and if it is a good enough fit to the data, it allows us to model a large data set approximately by a simple linear relationship. In the present experiment, the slope of this line will yield the coefficient of restitution.

Let our experimental data set be x_i and y_i , a collection of x and y values. Our first data point is $i = 1$, (x_1, y_1) , and the fifth is (x_5, y_5) . The trend line we seek can be described by

$$y_{\text{fit}} = mx + b \tag{4}$$

Once we determine the proper slope m and intercept b for our trendline, we can plug in a value for our explanatory variable x and the equation predicts a value for our dependent variable y_{fit} . If we choose to plug in the same x values as in our experiment, the corresponding values of y_{fit} will not necessarily match the data exactly for all points. Rather, the best fit trendline is the *the one that minimizes the overall differences between the trendline and the data*.

1.2 Calculating the best-fit trendline

Finding the line of best fit is not terribly different from calculating standard deviation, as we did in the previous experiment. The basic idea is that rather than calculating the standard deviation with respect to the mean (average) value of the data, we calculate the standard deviation with respect to an arbitrary *line* drawn through the data. Recall the definition of standard deviation:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \tag{5}$$

A descriptive version of the equation might go like this:

$$\sigma = \sqrt{\frac{1}{\text{num. points} - 1} \sum_{i=1}^{\text{num. points}} (\text{data point} - \text{average})^2} \quad (6)$$

If we want to calculate the deviation from a *line* instead of the average, we would do this:

$$(\text{deviation}) = \sqrt{\frac{1}{\text{num. points} - 1} \sum_{i=1}^{\text{num. points}} (\text{data point} - \text{line})^2} \quad (7)$$

Or, using our equation for the trend line,

$$(\text{deviation}) = \sqrt{\frac{1}{n - 1} \sum_{i=1}^n (x_i - mx - b)^2} \quad (8)$$

The best possible trend line is the one that *minimizes* this overall deviation. Performing the minimization requires the use of calculus, but the result is simple enough. How do we calculate the best-fit trendline? First, you will need a data table: your x and y values, their squares, and their products. Something like the table below; we have included two header columns to give you a better idea of what is going on.

Table 1: Data table

i	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	2	4.7	4	22.09	9.4
2	3	7.1	9	50.41	21.3
3	5	11.2	25	125.44	56.0
4	7	14.7	49	216.09	102.9
totals	17	37.7	87	414.03	189.6
average	4.25	9.425	-	-	-

The sums are easily performed in Excel; if you do not know how to do this, ask your neighbors or your TA. Now we have enough data tabulated to determine the line of best fit to our experimental data. Again, the derivation of how this works is beyond the scope of our discussion, but the

quantities calculated above are enough:

$$m = \text{trendline slope} = \frac{(\text{sum of } x \text{ times } y \text{ column}) - (\text{num. points}) (x \text{ average}) (y \text{ average})}{(\text{sum of } x^2 \text{ column}) - (\text{num. points}) (x \text{ average})^2}$$
$$m = \frac{\left(\sum_{i=1}^{i=n} x_i y_i\right) - n \bar{x} \bar{y}}{\left(\sum_{i=1}^{i=n} x_i^2\right) - n \bar{x}^2} \quad (9)$$

Here the sums are defined as last time; the one in the numerator is the total of the column $x_i y_i$, the one in the denominator is the total of the x_i^2 column. The averages of all x and y values are \bar{x} and \bar{y} , respectively, and n is the total number of data points. For our data, we have

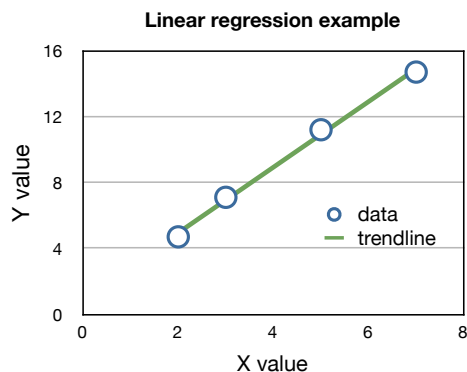
$$m = \frac{189.6 - 4 \times 4.25 \times 9.425}{87 - 4 \times 4.25^2} \approx 1.99 \quad (10)$$

The intercept is found more easily if we already have the slope:ⁱ

$$b = \text{trendline intercept} = \bar{y} - m \bar{x} = 9.425 - 1.99 \times 4.25 \approx 0.968 \quad (11)$$

Of course, these parameters m and b describing the trend line have uncertainties, which can be calculated. This is somewhat tedious, however, so we will leave that for a later experiment . . .

A neat fact is that trend line must pass through the dataset average point (\bar{x}, \bar{y}) . Here is a plot of our data and the calculated trendline:



It is quite good! But how good?

ⁱWe should note that one can also calculate the uncertainty or *confidence interval* for the slope and intercept, a measure of how reliable the extracted trendline parameters are. This is a little more difficult and nuanced, and we will not dwell on it. Be aware that the trendline parameters are merely *best fit* values – in any particular case, however, even the ‘best fit’ can be pretty terrible! The correlation coefficient, discussed below and reported by most programs like Excel or DataStudio, is a simple way to judge the quality of the trendline in describing your data.

1.3 Correlation coefficient

We can also define (but will not derive) a “goodness of fit” parameter r . If $r = 0$, there is no correlation between x and y – total randomness. If $r = -1$, the data are perfectly *negatively* correlated – a line with negative slope. If $r = +1$, the data is perfectly *positively* correlated – a line with positive slope. The closer $|r|$ is to 1, the better the correlation, while a small value of $|r|$ near zero indicates poor correlation. We would guess that our r should be positive, and close to 1 based on the plot above. Here’s how we calculate r , using the same data table we already generated:

$$\text{quality of fit} = r = \frac{n \left(\sum_{i=1}^{i=n} x_i y_i \right) - \left(\sum_{i=1}^{i=n} x_i \right) \left(\sum_{i=1}^{i=n} y_i \right)}{\sqrt{n \left(\sum_{i=1}^{i=n} x_i^2 \right) - \left(\sum_{i=1}^{i=n} x_i \right)^2} \sqrt{n \left(\sum_{i=1}^{i=n} y_i^2 \right) - \left(\sum_{i=1}^{i=n} y_i \right)^2}} \quad (12)$$

It is a bit fearsome-looking, and we will use Excel to calculate it automatically in general, but it just involves sums we’ve already done. In our case,

$$r = \frac{4 \times 189.6 - 17 \times 37.7}{\sqrt{4 \times 87 - 17^2} \sqrt{4 \times 414.03 - 37.7^2}} \approx 0.998 \quad (13)$$

Almost perfect positive correlation. Of course, real data is often not quite this good; if you find $r \approx 0.8 - 0.9$ consider it a good day. We should note that usually programs like Excel tell you r^2 rather than r ; the main operational difference is that $r^2 = 1$ for perfect correlation, and $r^2 = 0$ when there is no linear correlation.

Finally, both the trendline generation and quality of fit estimate can be done automatically for a given plot in Excel and DataStudio! Now that you’ve done it manually, it is no longer just another magic feature: it is based on solid mathematics. One note in passing, Excel will report r^2 rather than r for the quality of the fit. Values of r^2 extend from 0 to 1, with 0 being no correlation, and 1 being perfect correlation.

1.4 Linearization of a non-linear relationship

Based on our discussions of free-fall motion, we know that the vertical position versus time for an object dropped from rest at a height x_o is given by

$$x = \frac{1}{2} a_o t^2 + x_o = -\frac{1}{2} g t^2 + x_o \quad (14)$$

Here the second form explicitly assumes that $a_o = -g$ with the upward direction defined as positive

y. How can we make this into a linear relationship so we can make use of the regression analysis we just learned? We can start by noticing that we *would* have a straight line relationship if we had t instead of t^2 in the equation above. If that is the case, all we need to do is plot x versus t^2 instead of x versus t , or in other words we can change the dependent variable from t to t^2 . Mathematically, we can do that by making the replacement $u = t^2$:

$$x = \frac{1}{2}a_0u + x_o = -\frac{1}{2}gu + x_o \quad (15)$$

Now we have a straight line of slope $-g/2$ and intercept x_o , and we can perform our regression analysis on the substituted data $x(u) = y(t^2)$, plotting x versus t^2 and finding the appropriate trendline. For rather subtle statistical reasons, this procedure is *not* generally as accurate as if the data were linear in the first place, but it is typically a very good way to estimate the best-fit parameters for a non-linear relationship.

As another example, let's say we had an exponential relationship, like this:

$$x = x_o e^{-at} \quad (16)$$

In this case, we could linearize the equation by taking the natural logarithm of each side and re-arranging:

$$\ln x = -at + \ln x_o \quad (17)$$

Now if we plot $\ln y$ versus t we should get a straight line of slope $-a$ and intercept $\ln x_o$. Not every equation can be linearized in a simple way, but the general trick is one we will use quite often in introductory physics.

2 Preparatory Questions

Comment on these questions in your report.

1. Sketch what you expect the velocity versus time graph for a falling ball to look like.
2. Sketch the corresponding acceleration and position versus time graphs for a falling ball.
3. What do you think the influence of air resistance would be?
4. Make $x-y$ plots of the four data sets below and find the best fit trendline for each. What does this tell you about the limitations of purely statistical analysis?ⁱⁱ *There is an excel file downloadable at http://faculty.mint.ua.edu/~pleclair/ph101/labs/Lab2_falling-object/ with this data already entered. The preceding link should be clickable, by the way.*
5. Investigate the applet at <http://www.shodor.org/interactivate/activities/Regression/>

Data set 1		Data set 2		Data set 3		Data set 4	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.10	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	7	5.24
4	4.26	4	3.10	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Table 2: Four example data sets for analysis.

3 Supplies & Equipment

1. Computer with DataStudio, sensor interface, Excel
2. Rubber ball

4 Suggested Procedure

Now we'll compare your predictions to experiment. **To start the experiment:**

1. Mount the motion sensor on a stand a meter or so above the table if it is not already.

ⁱⁱThese four datasets are known as "Anscomb's quartet," specifically designed to have identical simple statistical properties. see http://en.wikipedia.org/wiki/Anscombe's_quartet for more information.

2. Orient the sensor so that the ultrasonic beam will be projected downward.ⁱⁱⁱ
3. Open Data Studio, select “Create Experiment.”
4. Select “Motion Sensor” under the list of sensors. Drag it to the interface box on the right side. When you are over the correct connections they will become highlighted; let go when this happens.
5. Plug the sensor connections (color coded) into the interface as now shown on the screen.
6. Double click on the motion sensor icon that is shown connected to the interface, select the Motion Sensor tab, and set the trigger rate (it may be called the ‘sample rate’) to 100.

Making plots and taking data:

1. On the left-hand side of the screen at the bottom, double click on ‘graph’ to create a graph.
2. Select the data you wish to have on the y axis (position, velocity, or acceleration). Choose position to start with, you can add a velocity graph later.
3. Alternatively, in the upper right side of the screen, there is a list of available data. Drag ‘position’ to the ‘graph’ icon on the lower left of the screen, and a position-time graph should appear.
4. Hit “start” and “stop” to collect position data. While the data collection is running, you can verify that the sensor is working by moving your hand in front of it. *Note that the sensor measures position relative to itself.*
5. Once data has been collected, you can drag data from different ‘runs’ onto an existing graph or create a new one. You can also drag velocity data onto an existing position-time plot, and velocity-time plot will show up on the same graph.
6. During any given run, position, velocity, and acceleration data are *all* collected. You do not need to do a separate run to record velocity and position, you can always plot both from a given run.
7. Zooming, scaling, and so on should be straightforward - play with the controls a bit!

Experiment

1. Record the motion of the ball as it is dropped from about 20 cm below the motion sensor onto the table.
2. Try to record the motion of the ball as it bounces at least three times. You may need to attempt this a few times. If the ball drifts out of the ultrasonic beam as it bounces, the sensor will not detect its motion.
3. What differences and similarities do you observe between the measured position and your predicted graph? Can you explain the differences?
4. Now we want to think about the velocity of this ball. You already sketched what you think the velocity versus time graph will look like. Compare your description with your neighbors^Ö,

ⁱⁱⁱNote also the beam width selector switch on the back - narrow or wide beam. It should be set to narrow beam.

- and try to reach a consensus. Sketch your consensus prediction. What is the mathematical relationship between the position and velocity graphs?
5. Double click on “Graph” under “Displays” and drag velocity (under Data) onto the position graph so that you have simultaneous plots of position and velocity. Click the appropriate button on the menu bar to align and match the horizontal axes of the graphs. Does your velocity graph look the way you expected? Explain any differences you see.
 6. You can also use the “slope tool” to find the slope of the position-time curve at any point. How does this slope compare to the measured velocity?
 7. You also predicted what you think the acceleration graph will look like. Compare with your neighbors, and try to reach a consensus. Make a new sketch if necessary when you reach a consensus.
 8. Drag acceleration over to the graph as you did with velocity, and see if your prediction was correct. Discuss any differences you observe.
 9. Let us describe these graphs mathematically. Look first at the standard equations for a body in free fall, to see how you might determine the acceleration of the ball from these equations. Select a time interval between bounces; highlight those data points.
 10. Fit the position data to a **quadratic equation** and determine the acceleration of gravity from your fit parameters. Do this by clicking on the small “Fit” button and choose “quadratic.”
 11. Likewise, fit the velocity data between bounces to a **linear** equation and determine the acceleration of gravity. Note that the standard deviation (uncertainty) for slope and intercept are reported for linear fits, as is the correlation coefficient r . report your acceleration as the slope plus or minus the standard deviation.
 12. Now find the average acceleration of the ball during this time interval from the acceleration graph. (To find the average of a set of data points, click on the “ Σ ” button and choose “mean.”) You can also display the standard deviation σ and number of points n . Report the acceleration with uncertainty, the latter determined by σ/\sqrt{n} .

5 Discussion Topics for Report

These are questions you should address in your report, in addition to the preparatory questions and the questions embedded in the experimental procedure above. You should also include printouts of the position, velocity, and acceleration graphs showing the fits (position and velocity) and averaging (acceleration).

- Are the accelerations determined from $x(t)$ and $v(t)$ of the same magnitude?
- Do they have the same direction?
- Which experimental data gives the best estimate for g ? Why?

- Can you *quantify* the uncertainty in your data?

6 Format of Report

Your report need not be formal, the format is largely up to you. Answer all the questions above, turn in the plots you were asked to make, and your overall conclusions. Address the discussion topics briefly.