# Chapter 22

# Day 17: Introduction to Regression + Rubber Band Activity

## Learning Objectives

Today's assignment is an introduction to linear regression, as presented through the lens of vector calculus (i.e. regression as an optimization problem). We will then use linear regression to estimate the parameters of a rubber band from measurements collected during the day activity. Your estimated parameters will be used next week as part of the Jungle Bridge module.

## 22.1 Overview

The overall goal of the Jungle Bridge module is for you to build a model that accurately predicts the shape of a hanging structure built from weights and rubber bands/string. To accomplish this, we need to characterize the behavior of the rubber bands used to form the structure. Specifically, we would like to describe how the force exerted by the rubber band, $F$, depends on its stretched length, $l$. Rubber bands exhibit spring-like behavior: as you stretch them, the force they exert increases.

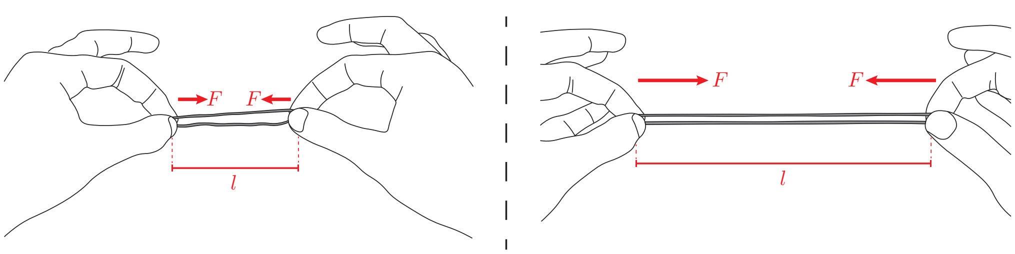


Figure 22.1: As the stretched length, $l$, of a rubber band increases, so does the force $F$ that it exerts at its end.

We usually use Hooke's law to model springs:

$$F = \begin{cases} k(l - l_0) & \text{for: } l \geq l_0 \\ 0 & \text{for: } 0 \leq l \leq l_0 \end{cases} \tag{22.1}$$

In other words, the force, $F$, scales linearly with the length, $l$. Hooke's law is governed by two parameters:

- The stiffness, $k$, which describes how a change in the stretched length of the rubber band, $\Delta l$, corresponds to a change in the output force, $\Delta F$:

$$\Delta F = k \Delta l \tag{22.2}$$

  This is the slope of the graph of $F$ vs. $l$.

- The natural length, $l_0$, which is the length for which the rubber band exerts zero force:

$$F\big|_{l=l_0} = 0 \tag{22.3}$$

  This is the x-intercept of the graph of $F$ vs. $l$.

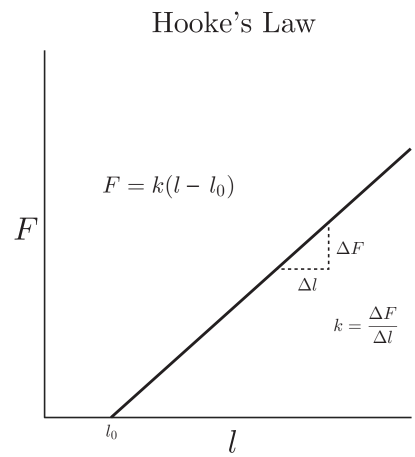


Figure 22.2: The stiffness, $k$, and natural length $l_0$ are the slope and x-intercept of the plot of $F$ vs. $l$.

Knowing $k$ and $l_0$ allows us to predict how a rubber band will behave as part of the hanging structure. The question then becomes, how do we measure the stiffness and natural length of a rubber band? Though we cannot measure the stiffness, $k$, directly, what we can do is measure the force that the rubber band exerts at different lengths, and then fit a straight line to the data. The stiffness and natural length can then be estimated as the slope and x-intercept of the fit line. The process of fitting a line to data is called simple linear regression which is a specific type of linear regression (which is itself a type of regression). As we'll soon find, regression can be thought of as a type of optimization problem, which can be elegantly described through the language of vector calculus.

## 22.2 Simple Linear Regression Overview (Brief)

Given a set of data points $(x_i, y_i), i \in [1, \ldots, n]$, we want to find the line, $y = mx + b$, that best fits the data:

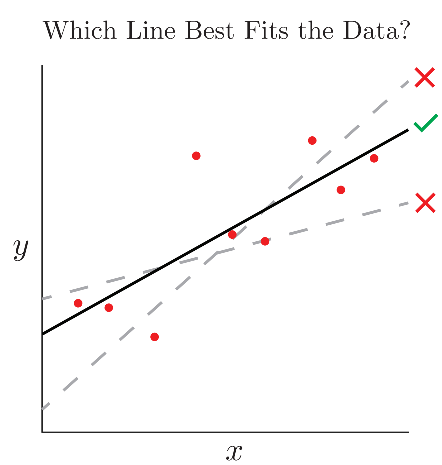


Figure 22.3: The goal of linear regression is to find the line that best fits the data.

This is the simple linear regression problem. In order to find this "line of best fit", we first need to quantify how well a line fits the data. For a given data point $(x_i, y_i)$, we can define the model error, $\Delta y_i$, as the difference between the value of $y_i$ predicted by the fit line, and the value that was actually measured for $y_i$:

$$\Delta y_i = (mx_i + b) - y_i \tag{22.4}$$

The larger the magnitude of $\Delta y_i$ gets, the larger the model error, which suggests a poor fit. As such, the line of best fit should minimize the model error across all data points. One way to evaluate this cumulative error across all data points is to take the sum of the squares of each of the errors:

$$E(m, b) = \sum_{i=1}^{n} (\Delta y_i)^2 = \sum_{i=1}^{n} (mx_i + b - y_i)^2 \tag{22.5}$$

If we can find a pair of values for $m$ and $b$ that minimize the cumulative error function, $E(m, b)$, then we will have found or line of best fit. Regression can be thought of as a type of optimization problem (a problem for which the goal is to find the best element of a set that either maximizes or minimizes some criteria).

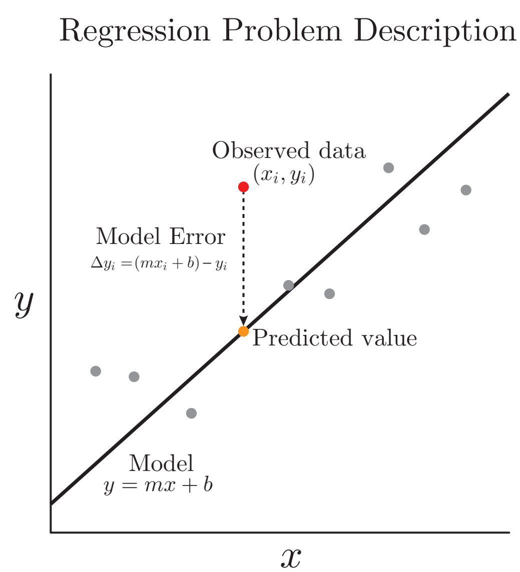


Figure 22.4: We can frame linear regression as an optimization problem, in which the goal is to minimize the cumulative difference between the model and observed data ($\Delta y_i$) across all data points $(x_i, y_i)$.

Since we want to find the best slope/intercept pair, $(m, b)$, that minimizes the total squared error across all observed data points, we can write the optimization problem down as follows:

$$\text{Find the value of } (m, b) \text{ that minimizes } E(m, b) = \sum_{i=1}^{n} (\Delta y_i)^2 \tag{22.6}$$

Or, in the more standard mathematical notation:

$$\underset{(m,b)}{\text{minimize}} \quad E(m, b) = \sum_{i=1}^{n} (mx_i + b - y_i)^2$$

Decision Variables　　　　Cost Function

An optimization problem has three main components:

- **Decision variables** (in this case $m$ and $b$): these are the variables that can be changed in the optimization in order to get a better/worse result.
- **Cost function** (in this case $E(m, b)$): a function of the decision variables (and other parameters) that we want to either minimize or maximize.

- **Constraints** (omitted...for now): Sometimes, there will be limitations that are placed on the decision variables. These constraints can come in a variety of flavors. An optimization problem with constraints is called a constrained optimization; while an optimization problem without constraints is called an unconstrained optimization. We will explore constrained optimization in a later section.

Since $E(m,b)$ is a function of two variables, the slope and intercept of the best fit line, $(m^*, b^*)$, correspond to the lowest point on the surface defined by $E(m,b)$:

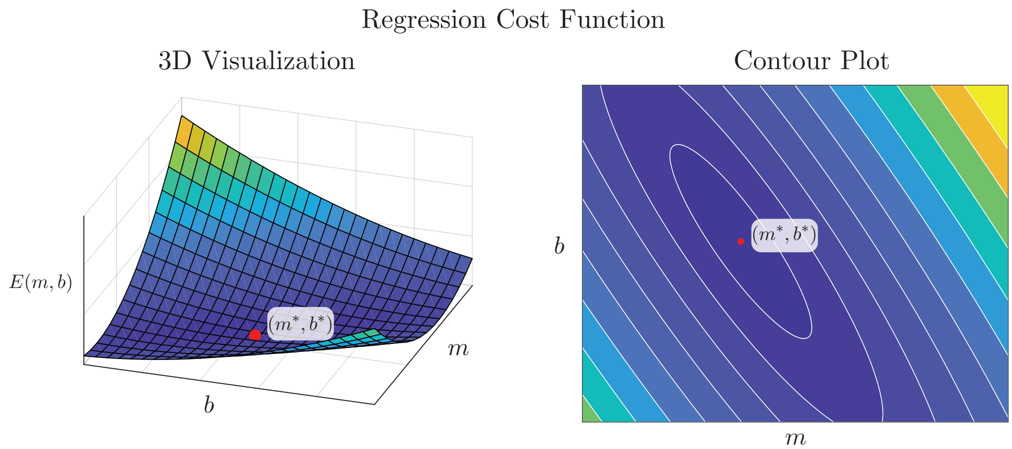


Figure 22.5: Linear regression chooses the slope and intercept values, $(m^*, b^*)$, that minimize the total squared error between the model and observed data, $E(m,b)$.

It can be shown that $(m^*, b^*)$ can be computed via the following formula:

$$\boxed{\begin{bmatrix} m^* \\ b^* \end{bmatrix} = \left(A^T A\right)^{-1} A^T Y} \tag{22.7}$$

Where $A$ and $Y$ are the following $n \times 2$ matrix and $n \times 1$ column vector of observations:

$$A = \begin{bmatrix} x_1, & 1 \\ x_2, & 1 \\ \vdots & \vdots \\ x_n, & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{22.8}$$

A proof of this equation 22.7 is provided in the appendix of this assignment. We strongly encourage you to go through this proof outside of class (however, please allocate your class-time to the completion of the data collection and analysis portions of today's activity).

## 22.3 Rubber Band Activity

### 22.3.1 Step 1: Making Nickel Weights

Your first task is to make weights out of nickels. We will be using these weights for both the rubber-band characterization and for later parts of the Jungle Bridge, so it's worthwhile to put in your best effort here. To make each weight, you will need the following items:

- A roll of scotch tape.
- 5-10 nickels.
- A medium-sized paperclip.
- A pair of scissors (to cut the tape with).
- A pair of pliers (to bend the paperclip).

Figure 22.6: The supplies you'll need for making weights out of nickels.

Begin by taping 5-10 nickels into a cylinder:

Figure 22.7: Tape nickels into a cylinder.

Bend the paperclip as illustrated below:

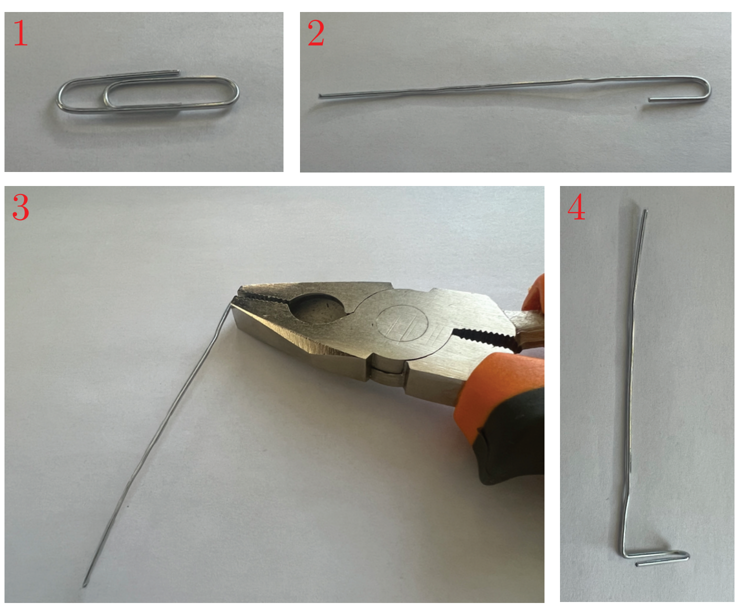


Figure 22.8: Bend paperclip as shown above.

Tape the bent paperclip to the **side** of the cylinder:

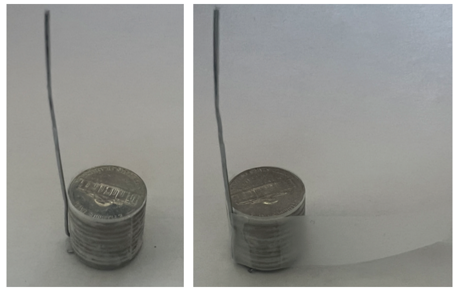

Figure 22.9: Tape the bent paperclip to the side of the nickel cylinder.

Tape the bent paperclip to the **bottom** of the cylinder:

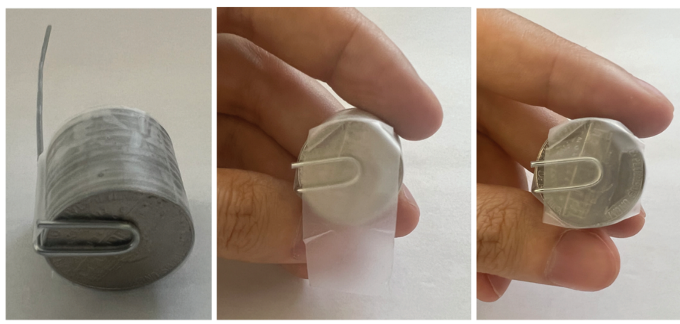

Figure 22.10: Tape the bent paperclip to the bottom of the nickel cylinder

Tape the side of the cylinder again so the edges from the bottom taping are secured to the cylinder:

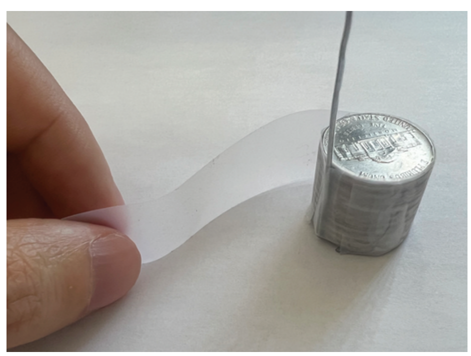

Figure 22.11: Tape the side of the cylinder again.

Use pliers to bend the top of the paperclip into a hook:

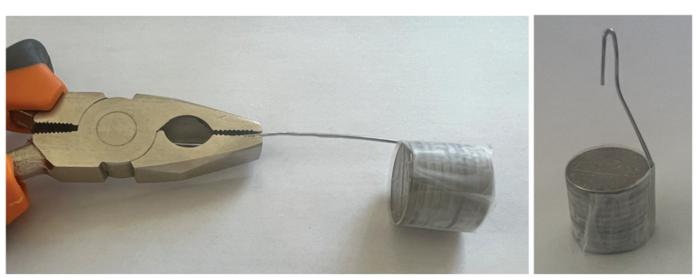

Figure 22.12: Use pliers to bend the top of the paperclip into a hook.

You will need to make 5-6 weights for Jungle Bridge module.

### 22.3.2 Data Collection

Your next task is to collect some data. You will be measuring how much a given rubber band stretches as a function of the amount of weight that is attached to it. To do this, you will need the following items:

- The rubber band that you want to characterize.
- 4-5 weights.
- A kitchen scale.
- A ruler.
- Paper labels (not displayed).

**Note:** It is extremely important that you keep track of the different rubber bands that you are characterizing. For this purpose, we have provided a set of paper labels and ziploc bags. Please label the rubber bands that you measure, and then store them in a bag so you don't lose them! We have also provided an excel template (on Canvas) that you can use for recording your data.

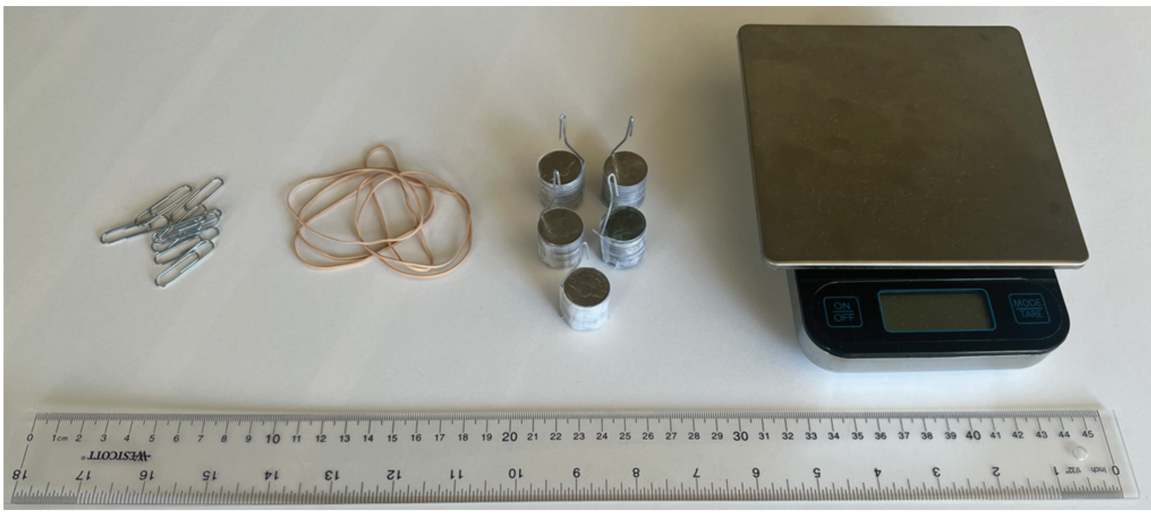

Figure 22.13: The supplies you'll need for measuring the stiffness and natural length of the rubber band.

Choose one or more weights (to attach to the rubber band). Use the kitchen scale to measure the total mass of these weights. Record this measurement in a table (the recommended unit is grams).

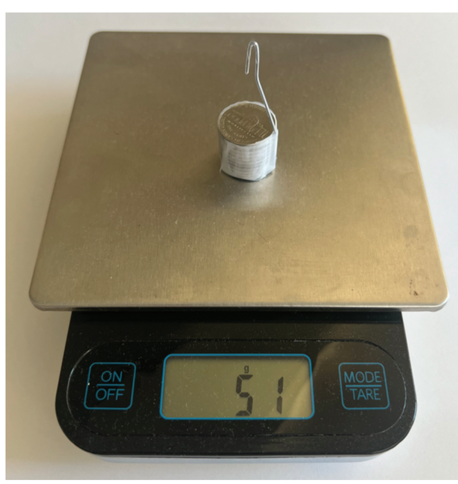


Figure 22.14: Use the kitchen scale to measure the mass of the weights to be attached to the rubber band.

Attach paperclips to the ends of the rubber band (as shown below). One of the paper clips will be used for holding the rubber band with your fingers. The other paper clip will serve as an attachment point for the weights. This will allow us to more accurately measure the stretched length of the rubber band w/o interference from our fingers (or the weights bouncing off the ruler).

Figure 22.15: Attach paperclips to the ends of the rubber band, one for holding with your fingers, the other as an attachment point for the weights.

Hang the weights from one of the paperclips. Pick up the other paperclip with your fingers until the weights are suspended in the air via. the rubber bands. Use a ruler to measure the stretched length of the rubber band. Record this measurement in a table (the recommended unit is centimeters).

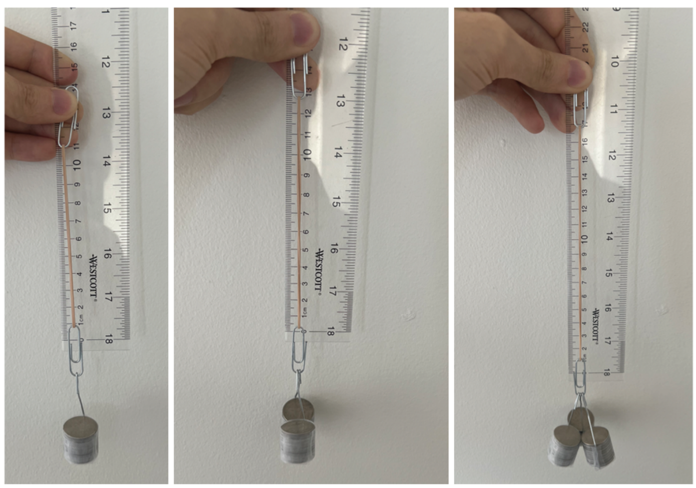

Figure 22.16: Measure the stretched length of the rubber band for the attached weight. Repeat this process for the same rubber band using different amounts of weight.

Repeat this process multiple time for the same rubber band using different amounts of weights (we recommend that you collect 3-5 data points per rubber band). We recommend that you collect data for 5-7 rubber bands . Make sure to label each rubber band that you measure so that you can keep track of them.

### 22.3.3 Estimating the Stiffness and Natural Length Using Regression

Now that we have collected the relevant data, our first task is to load our table into MATLAB. This can be accomplished using the `readtable` and `table2array` functions, as shown in the example below:

```
%This block of code demonstrates how to data from
%an excel file into MATLAB
function load_excel_example()
    %specify the path of the folder where the excel file is saved
    %make sure either the path string ends in \
    %or the file name string begins with \
    fpath = 'C:\Users\taylorott\Dropbox (Personal)\OrionTeachingMaterials\';
    %specify the file name of the excel file we want to load
    fname = 'RubberBandTemplate.xlsx';

    %use read table to load excel file into variable my_table
    %note that my_table is a MATLAB table
    %which is different from the usual MATLAB matrix data type
```

```
    my_table = readtable([fpath,fname]);

    %print entire table to the command line
    disp(my_table);

    %specify the row and column range of the numeric values
    %we want to extract from the table
    row_range = 3:8;
    col_range = 3:6;

    %print the block of the table specfied by row_range and col_range
    disp(my_table(row_range,col_range));

    %use table2array to convert desired portion of table into matrix
    %make sure that the specified cells only contain numeric values!
    data_mat = table2array(my_table(row_range,col_range));

    %display the matrix extracted from the excel file
    disp(data_mat);
end
```

We want to compute the line of best fit, where the independent variable is the stretched length of the rubber band, $l$, and the dependent variable is the force exerted by the rubber band, $F$:

$$F = ml + b \tag{22.9}$$

We can then use $m$ and $b$ to find $k$ and $l_0$ by equating the two linear relationships:

$$F = ml + b = k(l - l_0) = kl - kl_0 \tag{22.10}$$

Solving for $k$ and $l_0$, we get:

$$k = m, \qquad l_0 = \frac{-b}{m} \tag{22.11}$$

If we want to use equation 22.7 to compute the slope and y-intercept ($m$ and $b$) of the line of best fit, we must first use the measured data to construct the matrix $A$ and column vector $Y$, where:

$$A = \begin{bmatrix} l_1, & 1 \\ l_2, & 1 \\ \vdots & \vdots \\ l_n, & 1 \end{bmatrix}, \qquad Y = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{bmatrix} \tag{22.12}$$

Which is the goal of the following exercise:

### Exercise 22.1

In MATLAB, construct the corresponding $A$ matrix and $Y$ vector for a single rubber band. The `ones` function may be helpful for constructing $A$. Assuming that you recorded the masses of the weights in grams, you will need to compute the corresponding gravitational force in Newtons:

$$F = mg, \qquad g = 9.8 \text{ m/sec}^2 \tag{22.13}$$

Once $A$ and $Y$ have been constructed, we can use equation 22.7 to compute the slope and y-intercept ($m^*$ and $b^*$) of the line of best fit:

$$\begin{bmatrix} m^* \\ b^* \end{bmatrix} = \left(A^T A\right)^{-1} A^T Y \tag{22.14}$$

These values can then be converted into the stiffness and natural length, as previously described:

$$k = m, \qquad l_0 = \frac{-b}{m} \tag{22.15}$$

Let's do it in MATLAB!

**Exercise 22.2**

Use the values of $A$ and $Y$ that you constructed in MATLAB to compute this stiffness, $k$, and natural length, $l_0$ of the rubber band. **Note:** we encourage you to use the \ operator instead of the `inv` when computing $\left(A^T A\right)^{-1} A^T Y$. This is because \ uses Gaussian elimination, which is (numerically) much better behaved than the matrix inverse operation.

**Exercise 22.3**

Modify your code so that it computes the stiffness and natural length for each rubber band in the table. We recommend that you use a for loop to accomplish this.

## 22.4 Lab Report Deliverables

The primary deliverable for the Jungle Bridge module will be a lab report that documents your experiments and analysis. Today's rubber band analysis is an important part of this documentation. Please make sure to include the following in your lab report:

- For a single rubber band of your choice, include a plot comparing the line of best fit to the measured data. See regression overview for an example of such a plot.
- For the same rubber band, include either a contour plot or a 3D surface plot showing the cost function $E(m, b)$. See regression overview for an example of such a plot.
- Include a table of your measured data (for all rubber bands).
- Include a table showing the estimated values of $k$ and $l_0$ for each rubber band.
- Make sure to include your code with your final submission (don't put the code inside the lab report itself, however).
- Include descriptions of your experimental methodology and analysis. This should be written/summarized by you. Do NOT just copy and paste instructions from this assignment.
- Please refer to the style guide and the Jungle Bridge project document for further instructions.

## 22.5 Appendix: Regression Derivation

Given the cost function $E(m, b)$, which relates the slope and intercept of a fit line to the total squared error, we would like to find the optimal values, $(m^*, b^*)$ that minimize $E$:

$$\underset{(m,b)}{\text{minimize}} \quad E(m, b) = \sum_{i=1}^{n} (mx_i + b - y_i)^2 \tag{22.16}$$

To this end, it will be useful to define the column vectors $Y$, $\Delta Y$, $X$, and $\mathbf{1}$, where each column vector corresponds to a stack of values of $y_i$, $\Delta y_i$, $x_i$, and 1:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \Delta Y = \begin{bmatrix} \Delta y_1 \\ \Delta y_2 \\ \vdots \\ \Delta y_n \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{22.17}$$

Since the model error, $\Delta y_i$, is given by:

$$\Delta y_i = (mx_i + b) - y_i \tag{22.18}$$

We see that $\Delta Y$ can be expressed in terms of $X$, $\mathbf{1}$, and $Y$ as follows:

$$\Delta Y = \begin{bmatrix} \Delta y_1 \\ \Delta y_2 \\ \vdots \\ \Delta y_n \end{bmatrix} = \begin{bmatrix} (mx_1 + b) - y_1 \\ (mx_2 + b) - y_2 \\ \vdots \\ (mx_n + b) - y_n \end{bmatrix} = m \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = mX + b\mathbf{1} - Y \tag{22.19}$$

Since $E(m, b)$ is the sum of the model errors squared, it's equal to the Euclidean norm of $\Delta Y$ (squared)!

$$E(m, b) = \sum_{i=1}^{n} (\Delta y_i)^2 = ||\Delta Y||^2 \tag{22.20}$$

Substituting in our expression for $\Delta Y$, we get:

$$E(m, b) = ||mX + b\mathbf{1} - Y||^2 \tag{22.21}$$

As we learned previously, for $E(m, b)$ to be at a local minimum, its gradient must be zero:

$$\nabla E(m^*, b^*) = 0 \tag{22.22}$$

As such, we can find the optimal slope and intercept values, $(m^*, b^*)$, by computing $\nabla E$ and setting it to zero. Remember that the gradient of a function is equal to the vector of partial derivatives:

$$\nabla E(m, b) = \begin{bmatrix} \frac{\partial E}{\partial m} \\ \frac{\partial E}{\partial b} \end{bmatrix} \tag{22.23}$$

Since our cost function $E$ has the form $E(m, b) = ||V||^2$ where $V = (mX + b\mathbf{1} - Y)$, we can compute its partial derivatives (with respect to a variable $q$) as follows:

$$\frac{\partial}{\partial q}\left(||V||^2\right) = \frac{\partial}{\partial q}(V \cdot V) = V \cdot \frac{\partial V}{\partial q} + \frac{\partial V}{\partial q} \cdot V = 2V \cdot \frac{\partial V}{\partial q} \tag{22.24}$$

In the case of $E(m, b) = ||V||^2 = ||mX + b\mathbf{1} - Y||^2$, we get:

$$\frac{\partial E}{\partial m} = 2(mX + b\mathbf{1} - Y) \cdot \frac{\partial}{\partial m}(mX + b\mathbf{1} - Y) = 2(mX + b\mathbf{1} - Y) \cdot X = 2X^T(mX + b\mathbf{1} - Y) \tag{22.25}$$

$$\frac{\partial E}{\partial b} = 2(mX + b\mathbf{1} - Y) \cdot \frac{\partial}{\partial b}(mX + b\mathbf{1} - Y) = 2(mX + b\mathbf{1} - Y) \cdot \mathbf{1} = 2\mathbf{1}^T(mX + b\mathbf{1} - Y) \tag{22.26}$$

Plugging these expressions in for the gradient of $E$, we get:

$$\nabla E(m, b) = \begin{bmatrix} \frac{\partial E}{\partial m} \\ \frac{\partial E}{\partial b} \end{bmatrix} = 2 \begin{bmatrix} X^T(mX + b\mathbf{1} - Y) \\ \mathbf{1}^T(mX + b\mathbf{1} - Y) \end{bmatrix} \tag{22.27}$$

This can be rearranged into the following matrix product:

$$\nabla E(m, b) = 2[X \mid \mathbf{1}]^T(mX + b\mathbf{1} - Y) \tag{22.28}$$

At this point, it is convenient to define the $n \times 2$ matrix $A$ as follows:

$$A = [X \mid \mathbf{1}] = \begin{bmatrix} x_1, & 1 \\ x_2, & 1 \\ \vdots & \vdots \\ x_n, & 1 \end{bmatrix} \quad \rightarrow \quad mX + b\mathbf{1} = A \begin{bmatrix} m \\ b \end{bmatrix} \tag{22.29}$$

Substituting $A$ into our expression for $\nabla E$, we get:

$$\nabla E(m, b) = 2A^T\left(A\begin{bmatrix} m \\ b \end{bmatrix} - Y\right) = 2A^T A \begin{bmatrix} m \\ b \end{bmatrix} - 2A^T Y \tag{22.30}$$

Setting $\nabla E$ to zero (to solve for the local minimum), we get:

$$\nabla E(m^*, b^*) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \rightarrow \quad 2A^T A \begin{bmatrix} m^* \\ b^* \end{bmatrix} - 2A^T Y = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{22.31}$$

This equation is linear in $m^*$ and $b^*$, and its solution is given by:

$$\boxed{\begin{bmatrix} m^* \\ b^* \end{bmatrix} = \left(A^T A\right)^{-1} A^T Y}, \qquad A = \begin{bmatrix} x_1, & 1 \\ x_2, & 1 \\ \vdots & \vdots \\ x_n, & 1 \end{bmatrix}, \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{22.32}$$