Week 1: Notes for the instructor

Purpose and summary

The purpose of discussion this week is primarily to familiarize the students with R/R-studio and to set up why log-log and semi-log plots are useful in science. In the first problem on the discussion sheet, they will explore how to plot data and just generally get familiar with how R works. In the second problem, they will plot data that is badly scaled. That is, when plotted on linear axes, it is difficult to intuit the functional relationship. When plotted on log-log axes, the data are (approximately) linear. Note that, on the worksheet, it is never explicitly stated that the data should be plotted on log-log axes. This is so that when the instructor discusses log-log plots in lecture, these data can be used as an example.

Things to do in lecture

To ensure that students are prepared for section, we have found that it helps to:

- Announce that students must install R/R-studio on their computers/laptops and stress that R will be used each week in discussion and (if applicable) on homework.
- Announce that students must bring their laptops to section. While some students might not have laptops, the important thing is for each group in section to have at least one laptop.
- Assign an R-tutorial for the students to complete prior to section.

Materials

We have developed the following material to assist you:

1. Instructions for how to install R/R-studio (for Mac and PC).
2. An R/R-studio tutorial
Week 1: Pre-discussion assignments

- Install R/R-studio on your computers/laptops. Instructions for how to do this (for Mac and PC) can be found in separate documents.
- Bring your laptop to section.
- Complete the R/R-studio tutorial prior to section.
Problem 1. TRIAL AND ERROR

A zoologist is studying the relationship between weight \((w)\) and height \((h)\) for giant pandas. Based on initial measurements she forms a hypothesis for the functional relationship between weight and height for this animal,

\[ w = 0.2h^2. \]

She wants to confirm her hypothesis by comparing this relationship to measured data. The height of the panda varies between 0 and 3 meters. **Plot the functional relationship. To do this, type the R commands found in figure 1 in the console of R Studio (after the > prompt). There are 3 lines of code that need to be entered to generate a plot. Each row in the table is a line of code. Try all options and determine which one produces no errors.**

<table>
<thead>
<tr>
<th>1.</th>
<th>2.</th>
<th>3.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h = \text{seq(from}=0; \text{to}=3; \text{by}=0.01) )</td>
<td>( h = \text{seq(from}=0, \text{to}=3, \text{by}=0.01) )</td>
<td>( h = \text{seq(to}=0, \text{from}=3, \text{by}=0.01); )</td>
</tr>
<tr>
<td>( \text{w} = 0.2(h^2) )</td>
<td>( \text{w} = 0.2h^2 )</td>
<td>( \text{w} = 0.2*\text{h}^2 )</td>
</tr>
<tr>
<td>( \text{plot}(h, \text{w, xlab}=&quot;h (meters)&quot;}, \text{ylab}=&quot;w (kilograms)&quot;, \text{type}=&quot;l&quot;) )</td>
<td>( \text{plot}(h, \text{w, xlab}=&quot;h (meters)&quot;, \text{ylab}=&quot;w (kilograms)&quot;), \text{type}=&quot;l&quot;) )</td>
<td>( \text{plot}(w, \text{h, xlab}=&quot;h (meters)&quot;, \text{ylab}=&quot;w (kilograms)&quot;, \text{type}=&quot;l&quot;) )</td>
</tr>
</tbody>
</table>

**Figure 1: Code for Problem 1.**

Based on your findings, explain what each symbol in each line of code does.

<table>
<thead>
<tr>
<th>Table 1:</th>
<th>Measured height (meters)</th>
<th>Measured weight (kilograms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specimen 1</td>
<td>0.5</td>
<td>0.056</td>
</tr>
<tr>
<td>Specimen 2</td>
<td>1.0</td>
<td>0.44</td>
</tr>
<tr>
<td>Specimen 3</td>
<td>2.0</td>
<td>0.80</td>
</tr>
<tr>
<td>Specimen 4</td>
<td>2.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

On the same plot, the zoologist wants to plot the measured data contained in **table 15** below. To do this, type the following R commands in the console of R

```R
# Table 1: Measured height and weight
specimen_height <- c(0.5, 1.0, 2.0, 2.5)
specimen_weight <- c(0.056, 0.44, 0.80, 1.5)

# Plotting the measured data
plot(specimen_height, specimen_weight, xlab = "Measured height (meters)", ylab = "Measured weight (kilograms)", type = "l")
```
Studio (after the > prompt).

mh=c(0.5, 1.0, 2.0, 2.5)
mw=c(0.056, 0.44, 0.80, 1.5)
lines(mh,mw,type="p")

There are three main differences between the code we used previously to plot a function and this code for plotting scatter points. What are they?

Questions to think about:
Does the data agree with the zoologist’s hypothesis? Can you propose a method to quantify how “close” the predicted function is to the measured data?

Problem 2. ENERGY CONSUMPTION AND SIZE OF ANIMALS

Basal metabolic rate (BMR) is the minimal rate of energy expenditure per unit time by endothermic animals at rest. That is, the BMR is the amount of energy that an animal needs to keep the body functioning at rest through processes such as breathing, blood circulation, controlling body temperature, cell growth, brain and nerve function, and contraction of muscles. A human’s basal metabolic rate typically accounts for about 60 to 75% of the calories they burn every day.

Table 2:

<table>
<thead>
<tr>
<th>Approximate Body Weight, W (lbs)</th>
<th>Basal Metabolic Rate, BMR (kcal/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>elephant 10000</td>
<td>18000</td>
</tr>
<tr>
<td>horse 1000</td>
<td>3200</td>
</tr>
<tr>
<td>human 200</td>
<td>960</td>
</tr>
<tr>
<td>cat 8</td>
<td>86</td>
</tr>
<tr>
<td>rat 2</td>
<td>30</td>
</tr>
<tr>
<td>mouse .5</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 2 lists body weight (W) and basal metabolic rate (BMR) of several different mammals. Plot this data using R and try to determine the functional relationship (scaling law) between BMR and W.

To do this, type the following R commands in the console of R Studio (after the > prompt):

w = c(10000, 1000, 200, 8, 2, 0.5)
bmr = c(18000, 3200, 960, 86, 30, 10)
plot( w, bmr, xlab="W (lbs)", ylab="BMR (kcal/day)" )
... or for a slightly fancier plot, replace the last command with

```r
plot(mass, bmr, xlab="weight (lbs)", ylab="BMR(kcal/day)", cex=1.25, pch=21, bg="red", lwd=1)
```

(Try to figure out what each command does.)

**Questions to think about:**

Examine the graph of BMR and weight that you have just plot in R. Can you identify the relationship between an animal’s BMR and weight? Can you even distinguish all of the data points? What is preventing you from doing this? Can you think of a way to remedy the situation?
Problems for week 1:
Getting used to R, power and exponential functions

Problem 1. POWER VS. EXPONENTIAL FUNCTIONS
Using R/R-studio, plot the following functions on the same graph:

1. \( g = 2^r \)
2. \( g = r^2 \)
3. \( g = 2^r \)
4. \( g = e^r \)

Problem 2. POWER FUNCTIONS
Using R/R-studio, plot the following pairs of functions for \( L \geq 0 \):

1. \( W_1 = L^3 \) and \( W_2 = L^2 \)
2. \( W_1 = 1.8L^3 \) and \( W_2 = 0.3L^2 \)
3. \( W_1 = aL^3 \) and \( W_2 = bL^2 \), where \( a \) and \( b \) are parameters (i.e. constants).

Where do the graphs of each pair of functions intersect? Where is \( W_1 > W_2 \), and where is \( W_1 < W_2 \)?
Week 2: Notes for the instructor

NOTE: This is a particularly challenging discussion sheet, perhaps the most difficult one of the quarter. Part of what makes it so difficult is that the students have to develop a model, and are not explicitly told how to do so. The pre-discussion assignment can be a useful tool to get them thinking along the right lines.

Purpose and summary

The purpose of discussion this week is 1) to give the students experience in modeling; 2) to show the students why power laws are so important; and 3) to give the students experience with log-log plots.

The first question deals with how much water clings to an animal when it gets wet. The students will see that the weight of water scales as the surface area of the animal (proportional to the animal’s “radius” squared, if we model the animal as a sphere), while the weight of the animal scales as the volume of the animal (proportional to the animal’s radius cubed). The quantity $R$, which is the ratio of the weight of water to the weight of the animal, then scales as the animal’s mass to the negative $1/3$ power.

The second question deals with making a log-log plot with real data, eliminating outliers, and then interpreting that data to answer questions of biological interest.

It is possible to eliminate the second question, and only have question 1 on the worksheet. Problem 2, by itself, is probably too short for a complete worksheet.

Things to do in lecture

Do we want to include Tim’s lecture notes on Allometric scaling, or why aren’t there 10-foot-tall people?

Materials

Do we want to include Tim’s lecture notes on Allometric scaling, or why aren’t there 10-foot-tall people?
Week 2: Pre-discussion assignments

The discussion sheet is difficult. We have found it helps for the students to perform the following exercises prior to section, which help guide their thinking when completing the worksheet.

Pre-discussion questions

Use whatever resources you can (probably Google would be best), and your algebra skills to find the answers to these questions.

1. In the context of a data set, what is an outlier?

2. What is the surface area of a sphere? What’s the volume of a thin spherical shell, of width \( w \)?

3. What is the volume of a sphere?

4. What is the ratio of the volume of a thin spherical shell of width \( w \) to the volume of a sphere?
   (a) in terms of the sphere’s radius?
   (b) in terms of the volume of the sphere?
**Problem 1. A DANGER OF BEING SMALL: GETTING WET**

“A man coming out of a bath carries with him a film of water about one-fiftieth of an inch in thickness. This weighs roughly a pound. A wet mouse has to carry about its own weight of water. A wet fly has to lift many times its own weight and, as everyone knows, a fly once wetted by water or any other liquid is in a very serious position indeed.”


How do you think Haldane came up with these conclusions? Did he go out and weigh humans, mice, and flies before and after dipping them in water? Probably not. In fact, these statements are probably not that precise. The main point of Haldanes statement is that as you get smaller the more dangerous getting wet becomes. Let’s build an idealized mathematical model to see why this is.

a) Let $R$ denote the ratio between the mass of the water film clinging to an animal’s body after getting wet ($M_W$) and the dry animal’s body mass ($M_B$). Find an expression (i.e., a mathematical model) that approximates $R$ as a function of $M_B$. Sketch a graph of $R$ as a function of $M_B$.

- (Discussion question) Why is $R$ as a function of $M_B$ something we’d like to know?
- Construct your model by making simplifying assumptions, including:
  - assume that all animals are approximately the same shape, and model the body of an animal as a simple geometrical shape (i.e., a sphere)
  - assume that the mass of the water film on a wet animal is proportional to the animal’s body surface area, and use Haldane’s estimate of the thickness of the water film ($\ell = 0.05 \text{ cm}$).
  - assume that the density of an animal is the same as the density of water.

b) Use R/R-studio to compute $R$ for a human (60 kg), a cat (5 kg), a rat (0.25 kg), mouse (0.02 kg), a shrew (0.004 kg), a bee (0.0001 kg), a housefly (0.00002 kg), and
a mosquito (0.0000025 kg). Plot the data points corresponding to these animals on your graph $R$ vs $M_B$ from a).

c) What can you conclude from your mathematical analysis? Do your calculations agree with Haldane’s assertions? Do you see any issues with your/Haldane’s conclusions that getting wet poses serious dangers to small animals?

(Further discussion questions) Consider the assumptions that you made in coming up with your mathematical model. Are they reasonable assumptions? How could you make your model more precise? Do you think your basic conclusions would change significantly if you constructed a more precise (but probably more complicated) model?

Problem 2. Weight and wingspan of birds “Why did dodos go extinct?” and “What should the wingspan of a flying human be?”

Ornithologists have measured and cataloged the wingspans and weights of many different species of birds. The table below shows the wingspan $L$ for a bird of weight $W$.

a) Use R/R-Studio to make a scatter plot of the data. Make a semi-log plot, and a log-log plot. Are there any outliers in the data? Is there any justification for removing them from the set?

b) Find an exponential function model and a power function model for the data (with outliers removed).

c) Graph the models from b). Which fit appears better?

d) The dodo is a bird that has been extinct since the late 17th century*. It weighed about 45 pounds and had a wingspan of about 20 inches. Could the dodo fly? Why or why not? The dodo’s extinction correlates with the first European settlers arriving on the island of Mauritius (where the dodo lived). Why do you think the dodo went extinct?

e) Based on your model, what wingspan would we need to be flying humans?

[ see appropriate R commands on last page. ]

R CODE FOR PROBLEM 2 Type the uncommented lines of R-code (i.e., the lines without “#” at the beginning) in the console window.

# A. INPUT DATA AND PLOT

# data is listed in this order: # (Turkey vulture, Bald eagle, Great horned owl, Cooper’s hawk, Sandhill crane, Atlantic puffin, King penguin, # California condor, Common loon, Yellow warbler, Emu, Common grackle, Wood stork, Mallard, Dodo*)
<table>
<thead>
<tr>
<th>Bird</th>
<th>Avg. body weight (W, lb)</th>
<th>Avg. wingspan (L, in)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turkey vulture</td>
<td>4.40</td>
<td>69</td>
</tr>
<tr>
<td>Bald eagle</td>
<td>6.82</td>
<td>84</td>
</tr>
<tr>
<td>Great horned owl</td>
<td>3.08</td>
<td>44</td>
</tr>
<tr>
<td>Cooper’s hawk</td>
<td>1.03</td>
<td>28</td>
</tr>
<tr>
<td>Sandhill crane</td>
<td>9.02</td>
<td>79</td>
</tr>
<tr>
<td>Atlantic puffin</td>
<td>0.95</td>
<td>24</td>
</tr>
<tr>
<td>King penguin</td>
<td>29.0</td>
<td>28</td>
</tr>
<tr>
<td>California condor</td>
<td>17.8</td>
<td>109</td>
</tr>
<tr>
<td>Common loon</td>
<td>7.04</td>
<td>48</td>
</tr>
<tr>
<td>Yellow warbler</td>
<td>0.022</td>
<td>8</td>
</tr>
<tr>
<td>Emu</td>
<td>138</td>
<td>69</td>
</tr>
<tr>
<td>Common grackle</td>
<td>0.20</td>
<td>16</td>
</tr>
<tr>
<td>Wood stork</td>
<td>5.06</td>
<td>63</td>
</tr>
<tr>
<td>Mallard</td>
<td>2.42</td>
<td>35</td>
</tr>
<tr>
<td>Dodo*</td>
<td>45</td>
<td>20</td>
</tr>
</tbody>
</table>

# corresponding Weights of birds
\[ w = c(4.4, 6.82, 3.08, 1.03, 9.02, 0.95, 29, 17.8, 7.04, 0.022, 138, 0.2, 5.06, 2.42, 45) \]

# corresponding wingspan of birds
\[ s = c(69, 84, 44, 28, 79, 24, 28, 109, 48, 8, 69, 16, 63, 35, 20) \]

#plot data plot(w, s, xlab='weight (lbs)', ylab='wingspan (in)')

# B. TAKE LOG OF DATA (BOTH w and s) AND RE- PLOT

# take log_10 of weight and wingspan data wlog=log10(w) slog=log10(s)
#plot log-log data on a new graph plot(wlog, slog, xlab='log(weight)', ylab='log(wingspan)')

# What linear function \( y = a x + b \) would be a good fit to the log(wingspan) vs log(weight) data # except for outliers, where \( x = \log(\text{weight}) \) and \( y = \log(\text{wingspan}) \)?

# C. CREATE A LINEAR FUNCTION WITH SLOPE ‘‘a’’ and VERTICAL INTERCEPT ‘‘b’’
# AND PLOT THE # CORRESPONDING LINE ON A FIGURE THAT ALREADY EXISTS.

# fit your data with a line. These commands fit the unscaled data. You will need to modify # them to fit the log-log and semi-log transformed data
\[ P<-\text{lm}(s \sim w) \]
\[ \text{plot}(w, s); \text{lines}(w, \text{fitted}(P)) \]
Problem 1. CARBON DATING

Carbon dating is a powerful technique to determine when a particular piece of organic matter ceased to live. Here is a short explanation of how it works.

Carbon in the atmosphere occurs commonly in two isotopes (i.e. two alternate forms): Carbon-12 ($^{12}\text{C}$) and Carbon-14 ($^{14}\text{C}$). Carbon-14 is radioactive, meaning that it spontaneously decays into another element. In the atmosphere, it is constantly being formed and constantly decaying, so there is a constant fraction of Carbon-14 in the atmosphere (i.e. the ratio of $^{14}\text{C}$ to $^{12}\text{C}$ is a constant).

Plants incorporate this atmospheric carbon into the sugars and starches that make up their bodies, and animals, in turn, incorporate the carbon from plants into their bodies. However, when a plant or animal dies, it no longer exchanges carbon with the atmosphere, and the Carbon-14 in its body begins to decay. Eventually, after a very long time, the only carbon left is Carbon-12. For shorter times, if we know the fraction of Carbon-14 in the atmosphere, how quickly Carbon-14 decays and the fraction of Carbon-14 in a particular sample, we can determine when the sample died.

a) Carbon 14 has a half-life of 5,730 years, meaning that a sample of $^{14}\text{C}$ of initial mass $M(0) = m$ will have $M(5730) = m/2$. Given that $M(t)$ follows the equation

$$M(t) = M(0)e^{-\lambda t}$$

(1)

use the half-life to calculate $\lambda$.

(*Use R/R-studio to graph the function for different choices of $m$ and check your answer).

b) How many years would it take for the mass of the sample to drop to $m/4$?

(*Use R/R-studio to graph the function for different choices of $m$ and check your answer).

While snowshoeing in the Sierras over spring break, UC Davis student Brave X. Plorer got caught in a freak snowstorm. Seeking shelter from the elements, she and her dog Mittens entered a mountain cave. She lit her lantern and huddled back against the cave wall. To her surprise, the wall suddenly gave way and she tumbled, unhurt, into a large room. As she swung her lantern throughout the space, she noticed several unfamiliar animal skeletons and what appeared to be an intact spear.

Full of excitement, Brave resealed the room with a large rock, and curled up with Mittens for the night. In the morning, she hiked out and placed a phone call to the eminent Davis Professor R. Keyologist to explain her find. Returning later with the proper equipment, they retrieved the spear whose shaft was made of wood, preserved by the dry, sheltered environment of the sealed cave.

Back at Davis, they took a 1mg sample of the spear and, using accelerator mass spectrometry, determine that the ratio of $^{14}\text{C}$ to $^{12}\text{C}$ is 0.1 parts per trillion. Suppose that the ratio in the atmosphere is 1 part per trillion. There is therefore one tenth of the initial mass of $^{14}\text{C}$ in their sample.
c) How old is the spear? (round your answer to the nearest 1,000 years).

d) Does it pre-date the Clovis culture of approximately 11,540 years* ago? Such a finding would be very exciting and could make Brave X. Plorer famous.

**Problem 2. CHEMICAL BOND RUPTURE UNDER FORCE**

The interaction of biological molecules is critical to life. It is important, for example, for a white blood cell to be able to attach to the wall of a blood vessel in order to fight an infection. To do so, molecules on the surface of the white blood cell bind to molecules on the blood vessel. The white blood cell is often being pulled along by the flowing blood, so these chemical bonds experience mechanical forces.

One can measure the lifetime of chemical bonds under force. Here is a brief description of how it’s done (you don’t have to know this to solve the problem; it’s just in case you’re interested). This optional section starts at the line (above) and ends at the line (below).

---

Figure 2: Measuring bond lifetimes under force with an Atomic Force Microscope (AFM).

Bond lifetime can be measured with a device called an Atomic Force Microscope (AFM). A sketch of this apparatus is shown in Fig. 2, along with a sketch of some data.

To begin with, one isolates two molecules that bind to each other – say Molecule 1 and Molecule 2. Molecule 1 is attached firmly to the surface of a microscope slide. Then, Molecule 2 is attached to
the end of a small cantilever beam. They are brought in close proximity and a bond forms (panel 1, in the upper right of Fig. 2). A laser is aimed at the cantilever beam, so that it reflects off it and hits the center of a detector.

The surface is then moved away from the cantilever beam. This applies a force on the bond, causing the cantilever beam to bend. As the beam bends, the laser reflects to a different position on the detector. This deflection can be used to calculate the force on the bond (panel 2, in the upper center of Fig. 2).

Finally, the bond breaks. This cantilever beam relaxes and the laser reflection returns to the center of the detector (panel 3, in the upper right of Fig. 2). An experimental measurement looks like the trace on the bottom of Fig. 2. By moving the cantilever different distances, one can vary the force on the bond ($F$) and then measure the bond lifetime ($T$).

Forces affect how long a chemical bond lasts. Intuitively, one might expect that if you pull on a chemical bond, you shorten the lifetime of the bond. This is frequently the case.

There is a theory, called Bell’s Approximation, that predicts that bond lifetime ($T$) should depend on force ($F$) according to the equation

$$T = T_0 e^{-aF} \tag{2}$$

where $T_0$ and $a$ are constants that depend on the molecules that form the chemical bond.

a) Here are some measurements of bond lifetime at different forces (see Table 1). These measurements were made for a bond that forms between a molecule found on the surface of a white blood cell and a molecule found on blood vessels (Marshall et al., *Nature*: 423, 190–193. 2003.). Assuming that Bell’s Approximation is correct (Eq. 2), use R/R-studio to make a plot where the data fall along a (roughly) straight line.

<table>
<thead>
<tr>
<th>Force ($F$)</th>
<th>Bond Lifetime ($T$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>1.7</td>
</tr>
<tr>
<td>13</td>
<td>1.1</td>
</tr>
<tr>
<td>17</td>
<td>0.47</td>
</tr>
<tr>
<td>22</td>
<td>0.29</td>
</tr>
<tr>
<td>27</td>
<td>0.23</td>
</tr>
<tr>
<td>34</td>
<td>0.11</td>
</tr>
<tr>
<td>42</td>
<td>0.12</td>
</tr>
<tr>
<td>55</td>
<td>0.024</td>
</tr>
</tbody>
</table>

Table 3: Bond lifetime for different forces from Marshall et al. 2003.

b) Use R/R-studio to fit a line to this data (see code below, for an example). Using this line, estimate the parameters $T_0$ and $a$ in Eq. 2.
Here’s how to fit a line to data, using R/R-studio. Suppose I have one set of measurements, \( A = 1,2,4,5,7,11 \) and, for each of these measurements, a corresponding measurement \( B = 4.00,8.11,16.35,20.03,28.38,44.21 \). Suppose that \( A \) is our independent variable, and \( B \) our dependent variable. Then, we could write the following code to fit a straight line to \( B(A) \):

\[
\begin{align*}
> A &\leftarrow c(1,2,4,5,7,11) \\
> B &\leftarrow c(4.00,8.11,16.35,20.03,28.38,44.21) \\
> \text{lm}(B \sim A) \\
> \text{plot}(A,B); \text{lines}(A,0.07848+4.02030*A)
\end{align*}
\]

In this code, the first two lines define the variables \( A \) and \( B \). The third line fits a straight line, with the dependent variable listed first, before the tilde, and the independent variable listed second, after the tilde. For these particular data points, the function \text{lm} \ spits out an intercept of 0.07848 and a slope of 4.02030. The fourth line plots the data, along with the best fit line.

c) One big result of the paper by Marshall et al. is that, at small forces, these bonds last longer in the presence of force than in its absence (this is called a “catch” bond). However, at larger forces, bond lifetime decreases as force increases (this is called a “slip” bond). For the forces in Table 1, is this a catch or a slip bond? What, if anything, would be different about the parameters \( a \) and/or \( T_0 \) in Eq. 2 for a slip bond vs. a catch bond? Explain.
Problem 3. OLYMPIC WEIGHT LIFTING RECORDS

The following figure shows the world record (in lbs) in the clean and jerk, an olympic weight-lifting event. These data are shown as a function of weight class (in lbs).

As you can see, the data are well-fit by a straight line. The slope of that straight line is 0.574, and when $\log_{10}(\text{bodyweight(lbs)}) = 0$, it predicts $\log_{10}(\text{clean-jerk(lbs)}) = 1.382$.

What is the relationship between body weight, in lbs, and world record clean and jerk, in lbs?

To get full credit, your answer must (1) define variables for the two quantities, (2) give a function that outputs world record clean and jerk, in lbs, for any body weight, lbs, and (3) not contain any logarithms.
Problem 4. RATING STIMULUS INTENSITY, STEVENS’S LAW

The following plot shows data from an experiment where subjects were asked to rate the brightness of a light. The subjects’ rating, \( R \), is plotted as a function of the brightness of the light \( B \). Note that the vertical axis shows \( \log_{10}(R) \), and that the horizontal axis shows \( 10 \log_{10}(B/B_0) \), where \( B_0 \) is a reference intensity. The best-fit line to the data (pictured) has a slope of 0.0299 and a vertical intercept of \(-2.15\).

What is \( R(B) \)? (Note, your answer will include \( B_0 \), the reference intensity).

Figure 4: Data from Stevens, 1957
Problem 5. HEART RATE OF MAMMALS

Below, the resting heart rate \( r \) and body mass \( M \) of several mammals are shown on a log-log plot. The data are reasonably fit by a straight line with slope \(-1/4\) and vertical intercept 3.07.

\[
\log_{10}(r) = \log_{10}(M)^{-1/4} + 3.07
\]

a) Is the function that relates body mass \( M \) to heart rate \( r \), which we call \( r = f(M) \), an exponential function or a power law?

b) What is \( f(M) \)?
Week 3: Notes for the instructor

NOTE: The data discussed in this worksheet are revisited in worksheet 8.

Purpose and summary

The purpose of this worksheet is to 1) demonstrate a real-world example of exponential population growth; 2) show how semi-log plots are used; 3) analyze data with a discrete time model; 4) use R/R-studio to generate a sequence from a recursion; and 5) point out some subtleties regarding recursions, discrete time models and the connection to continuous time models.

The worksheet has a single problem, which discusses an endangered bird, the whooping crane. Driven nearly to extinction by the mid 1930s, the population has slowly rebounded over the past 80 years, and exhibits nearly exponential growth. In the worksheet, the students plot real data on a semi-log and log-log plot (using R) and observe that it is linear on the semi-log plot. Using R, they fit a line to the data on the semi-log plot, and generate a continuous time model, $N(t)$. They then use this model to generate a sequence, $P_t$, that predicts the number of cranes each decade – the same interval over which the data are measured. They then find a recursion for the sequence, $P_{t+1} = rP_t$, and compare this growth per decade, $r$, to the growth per year, $c$, in the continuous time model $N(t) = N_0c^t$. Since the former is per decade and the latter is per year, the relationship between the two is $c = r^{10}$.

Things to do in lecture

It is our hope that this worksheet fits well with the discussion of Chapter 2 in Neuhauser.

Materials

Because the worksheet contains some potential pitfalls for the students, we have provided some notes for TAs/LAs.
Problem 1. Conservation of whooping cranes  Whooping Cranes are large, heron-like birds that live in the west/mid-west of America (Alberta and Wisconsin) in the summer, and migrate south (Texas and Florida, respectively) in the winter. A combination of hunting and habitat loss nearly drove them to extinction in the 1930s. In 1967, the Whooping Crane was declared endangered, and it has been the subject of intense conservation efforts ever since. The conservation effort has had limited success due, in part, to the longevity of the birds (they can live to be over 25 years), and their long, difficult migration.

Below is a table showing the number of wild Whooping Cranes each decade from 1940 to 2010. It comes from a paper by Butler, Harris and Strobel, published in the journal Biological Conservation in 2013.

<table>
<thead>
<tr>
<th>Year</th>
<th>time since start (years)</th>
<th>Number of Cranes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>1950</td>
<td>10</td>
<td>32</td>
</tr>
<tr>
<td>1960</td>
<td>20</td>
<td>37</td>
</tr>
<tr>
<td>1970</td>
<td>30</td>
<td>58</td>
</tr>
<tr>
<td>1980</td>
<td>40</td>
<td>79</td>
</tr>
<tr>
<td>1990</td>
<td>50</td>
<td>146</td>
</tr>
<tr>
<td>2000</td>
<td>60</td>
<td>180</td>
</tr>
<tr>
<td>2010</td>
<td>70</td>
<td>283</td>
</tr>
</tbody>
</table>

a) Plot the raw data with the following commands:

\[
t = \text{seq}(\text{from = 0, to = 70, by = 10})
\]
\[
N = \text{c}(26, 32, 37, 58, 79, 146, 180, 283)
\]
\[
\text{plot}(t, N)
\]

Make a guess: is the data linear, power law or exponential? To confirm your intuition we will look at both possibilities by making a log-log plot (part b) as well as a log-linear plot (part c) and see which has a better fit to a line.
b) Type in the following lines of code to plot the data on a log-log scale then fit a line to that data

```r
t=seq(from = 10, to = 70, by = 10)
logt = log(t, base=10)
N=c(32,37,58,79,146,180,283)
G<-lm(log(N, base = 10) ~ logt)
plot(logt, log(N, base = 10))
lines(logt, fitted(G))
```

Does that look like a good fit? We will next try a log-linear and then decide which was a better fit.

c) Type in the following lines of code to plot the data as a straight line and then fit a line to that data

```r
t=seq(from = 0, to = 70, by = 10)
N=c(26,32,37,58,79,146,180,283)
F<-lm(log(N, base = 10) ~ t)
plot(t, log(N, base = 10))
lines(t, fitted(F))
```

d) Go back to your guess and look at your log-log and log-linear fits to lines. What is the functional relationship between $N$ and $t$ (i.e. power law, exponential, etc.)?

To get the equation for the best-fit line, type in “F” or “G” at the prompt in R-studio. If the linear fit is of the form $y = mx + b$, then the coefficient called “intercept” is $b$ and the other coefficient is the slope.

e) Based on your previous answer, find an equation for $N(t)$, which is a model for the number of whooping cranes as a function of time (where time is measured in years since 1940).

f) In part (e), your answer should be of the form $N(t) = ac^t$, or $N(t) = at^c$ where $a$ and $c$ are constants. You only have to find the model for one (power law OR exponential depending on your answer to (d)). What do the parameters $a$ and $c$ represent?

g) Your equation for $N(t)$ (from part e) is a model. It makes some assumptions, and we could test these assumptions. A critical assumption of the model is that the growth rate, $c$, does not vary with time. Therefore, even though we don’t have data for the years between our 10 year measurements, the model predicts them (i.e. the model predicts there were 247 cranes in 2009). In fact, the model even predicts the number of cranes down to the minute. Discuss and answer the following (in com-
plete sentences): Do you think the model is reliable in its predictions of the number of cranes over very short intervals? Are there reasons to suspect that growth rate might not be a constant over some time intervals? Hint: whooping cranes have a very specific breeding season, and migrate from Canada down to Texas.

h) The number of cranes at **10 year intervals** forms a sequence, \( N_0 = 26, N_1 = 32, N_2 = 37 \), etc. Use your model from part e (i.e. your equation for \( N(t) \)) to generate a sequence of the predicted number of cranes, \( P_0, P_1, \) etc., up to \( P_7 \).

i) Discuss and answer the following (in complete sentences): Are the predicted values close? Do you think you could do better? Where is the prediction the best? Where is it the worst? Could you predict the number of cranes 10 years in the future? 100 years? Would you have confidence in these predictions?

j) Find a recursion, of the form \( P_{i+1} = rP_i \) (where \( r \) is a constant that you must determine) that describes the predicted number of cranes. Don’t forget that your prediction in (e) assumes that time is measured in **years** while the data \( P_0, P_1, \) etc., up to \( P_7 \) is given in **10 year increments**. How does that affect your value for \( r \)? What does the constant \( r \) represent? How does it relate to the parameter \( c \) in your answers to parts (e) and (f)?

k) Check your answer to part (e) with a “for” loop in R-studio. Here’s an example (with made up numbers) for how it works. Pretend my recursion from part j) is \( P_{i+1} = 2P_i \), and suppose that \( P_0 = 10 \). Then, the following code will generate the sequence \( P_i \), stored in the variable \( P \). Note that the first entry of \( P \) will be \( P_0 \), the second \( P_1 \) etc.

```r
P=10
for(i in 1:7){P[i+1]=2*P[i]}
P
```

Bonus questions: In what year will there be 1,000 cranes? How about 10,000 – the estimated number of cranes that existed before humans arrived in America?
MAT 17A - DISCUSSION #3
Notes for TA’s and LA’s

a)-c) We want the students to investigate what makes a better fit, log-log or log-linear. Since the
time data starts at $t = 0$ we cannot use this for the log-log plot so we will use only the data starting
at (10, 32). This is a somewhat subtle point that we do not want to make a big deal of on the sheet
because it would be a distraction but if the students ask, that is what is going on in part b).

d) The answer is log-linear and exponential, if the students pick log-log they will not be able to go
on to do the recursion, you should make sure everyone is on the same page with this part.

e) The answer here is $N(t) = 21.4(1.035)^t$ where $t$ is measured in years since 1940.

f) Again here they should have an exponential, how do 21 and 1.03 relate to the slope and intercept
from the fit?

g) This discussion point should be answered in paragraph form. The students should talk about it
in discussion and write this up.

h) The predicted sequence is

<table>
<thead>
<tr>
<th>$P_0$</th>
<th>21.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>30.5</td>
</tr>
<tr>
<td>$P_2$</td>
<td>43.5</td>
</tr>
<tr>
<td>$P_3$</td>
<td>62.1</td>
</tr>
<tr>
<td>$P_4$</td>
<td>88.4</td>
</tr>
<tr>
<td>$P_5$</td>
<td>126.1</td>
</tr>
<tr>
<td>$P_6$</td>
<td>179.7</td>
</tr>
<tr>
<td>$P_7$</td>
<td>256.1</td>
</tr>
</tbody>
</table>

j) The recursion for this sequence is $N_{i+1} = 1.4253 N_i$, with $N_0 = 21.4$ Note that the time intervals
here are in 10 year increments so this will mean that $1.4253 = (1.035)^{10}$, this change of variables
may be confusing for the students, be sure to think it through.

k) The for loop does not have the right parameters, we want the students to figure that out, make
sure they see that it is just an example.
Problems for week 3: recursions, discrete time models

Problem 1. FITTING A MODEL TO DATA

As you may know, hemoglobin is a molecule in your blood that carries oxygen. It is a tetramer, meaning that each hemoglobin molecule consists of four identical protein molecules each of which binds one oxygen. Early studies of hemoglobin laid the groundwork for the field of biochemistry.

One researcher who studied hemoglobin was A.V. Hill. Hill was an extremely influential scientist: he won the Nobel Prize in physiology in 1922, made fundamental contributions to muscle physiology and developed a theory for enzymes that is as important as the Michaelis-Menten equation (which we saw in class). This problem relates to Hill’s work on hemoglobin in 1910 that led to his theory for enzymes.

At the time, it was assumed that Hemoglobin was a monomer, a single large molecule that binds a single molecule of oxygen (of course, now we know that hemoglobin occurs in the body as a tetramer). Hill suspected that under some experimental conditions, Hemoglobin was actually a dimer (two connected molecules, each binding a single oxygen molecule) or even a mixture of dimers and monomers. He therefore came up with a mathematical model that could describe oxygen binding to a monomer (Model 1), oxygen binding to a dimer (Model 2) and oxygen binding to a mixture of monomers and dimers (Model 3). These equations are given below, where each equation relates the amount of oxygen \( x \) to the saturation of hemoglobin \( y \):

Model 1: hemoglobin occurs only as monomers

\[
y = 100 \frac{0.0621x}{1 + 0.0621x}
\]  

(3)

Model 2: hemoglobin occurs only as dimers

\[
y = 100 \frac{0.00192x^2}{1 + 0.00192x^2}
\]  

(4)

Model 3: hemoglobin occurs as a mixture of monomers and dimers

\[
y = 82.6 \frac{0.00150x^2}{1 + 0.00150x^2} + 17.4 \frac{0.202x}{1 + 0.202x}
\]  

(5)
Hill then compared the models to measurements from Barcroft and Camis, given below (in modified form):

<table>
<thead>
<tr>
<th>Amount of oxygen ($x$)</th>
<th>% Saturated Hemoglobin ($y$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>22.2</td>
</tr>
<tr>
<td>20</td>
<td>45.6</td>
</tr>
<tr>
<td>30</td>
<td>60.7</td>
</tr>
<tr>
<td>40</td>
<td>75.3</td>
</tr>
<tr>
<td>50</td>
<td>83.6</td>
</tr>
<tr>
<td>60</td>
<td>80.2</td>
</tr>
<tr>
<td>70</td>
<td>91.8</td>
</tr>
<tr>
<td>80</td>
<td>89.1</td>
</tr>
<tr>
<td>90</td>
<td>94.9</td>
</tr>
<tr>
<td>100</td>
<td>94.2</td>
</tr>
</tbody>
</table>

Table 6: Data on hemoglobin saturation (Barcroft and Camis, 1909).

Based on these measurements, you will decide whether, in these experiments, hemoglobin is a monomer, dimer or mixture of monomers and dimers.

Note that we can consider the measurements in the table above to be a sequence $M_0 = 0, M_1 = 22.2, M_3 = 45.6, \ldots, M_{10} = 94.2$.

a) Use each model to generate a sequence of predicted measurements, $Y_i$. In particular, each model gives you a functional relationship between $y$ and $x$. Using this functional relationship, find the predicted percentage of saturated hemoglobin at $i\Delta x$, where $i = 0, 1, 2, 3, \ldots, 10$ and $\Delta x = 10$.

b) To determine how well the model fits the data, we’ll calculate the sum of the squared error for each model. To do so, generate a new sequence $E_1, E_2, E_3, \ldots, E_{10}$, where $E_i$ is the squared difference between the predicted measurement $Y_i$ and the actual measurement $M_i$ ($E_i = (Y_i - M_i)^2$).

c) The sum of the squared error, $S$, is the sum of all $E_i$’s. That is, $S = E_0 + E_1 + E_2 + \cdots + E_{10}$. Determine the sum of the squared error for each model.

d) The model with the smallest sum of squared error fits the data the best. Which one is it? Based on this result, in the experiments of Barcroft and Camis, was hemoglobin a monomer (Model 1), dimer (Model 2) or mixture of monomers and dimers (Model 3)?
Note that it might save you time and effort to complete these tasks using R/R-studio. Here are some potentially useful pieces of code:

1. Suppose I want to generate a sequence following the formula \( N_i = \frac{4i^2}{1 + i} \), and \( i \) going between 1 and 11. The following code will accomplish this task:

   ```r
   N = rep(0, 11)
   for(i in 1:11) {N[i]=4*i^2/(i+1)}
   ```

2. Suppose I ran the previous code, and want to define another sequence \( M \), which contains the elements of the data in Table 1. Then, I want to find yet another sequence \( E_i \) that is the error. The following code will do that:

   ```r
   M = c(0, 22.2, 45.6, 60.7, 75.3, 83.6, 80.2, 91.8, 89.1, 94.9, 94.2)
   E = (N - M)^2
   ```

3. Suppose I ran the previous code, and want the sum of the sequence \( E_i \). The following command will do that:

   ```r
   sum(E)
   ```
Week 4: Notes for the instructor

Purpose and summary

The purpose of this worksheet is to make the concept of differentiation intuitive. That is, we all know what it means for an object to be falling fast or slow, but the challenge (and the aim of this worksheet) is to connect that intuition with the formal definition of the derivative of the object’s position with respect to time.

The worksheet has two problems, the second one can easily be cut, should time be a concern. The first problem shows a very short movie of a penny falling. Along with that movie, there are six still frames from the movie. The students will estimate the change in the penny’s position in between the six frames and then divide by the time between frames to estimate the penny’s velocity. Since the penny is falling with little air resistance, its velocity increases (roughly) linearly with time. The slope of this line is the gravitational constant, which the students will then estimate by fitting their data with a straight line.

The second problem has the students determine a reasonable domain for a mathematical model, calculate the derivative of a quadratic function (using the limit definition) and then observe that the derivative is zero at a maximum, positive to the left, and negative to the right. These observations are meant to foreshadow the discussion of maxima and minima later in the course.

Things to do in lecture

This worksheet corresponds to chapter 4 in Neuhauser, but problem 1 can come before chapter 4 is discussed in lecture.

Materials

For problem 1 of this worksheet, there is a video (Penny_Drop.mov), which shows a penny falling in front of a piece of notebook paper. There are also six still frames from the movie (Frame_1.png, Frame_2.png, Frame_3.png, Frame_4.png, Frame_5.png, and Frame_6.png). The students need to have access to these materials during section in order to complete the worksheet.
Week 4: Pre-discussion assignments
MAT 17A - DISCUSSION #4

Problem 1. The falling penny

The concept of how a function’s output changes as its input changes is central to many branches of science. 17A instructor Sam Walcott made a video of a falling penny in front of a piece of lined paper. Watch the video (Penny_Drop.mov); it’s very short. Let’s calculate the penny’s velocity based on its change in position. The velocity is the change in the penny’s position ($y$) with respect to time ($t$), and it is written as follows

\[ v = \lim_{\Delta t \to 0} \frac{y(t + \Delta t) - y(t)}{\Delta t} = \frac{dy}{dt} \]

We will use the data from the movie to approximate the velocity. The movie goes very fast, and so to help with the calculation you are given six still frames (Frame_1.png, Frame_2.png, Frame_3.png, Frame_4.png, Frame_5.png, and Frame_6.png) equally spaced 1/23 seconds apart in time.

a) Fill out the table below to record the penny’s position as a function of time. The movie is recorded at 23 frames per second, and the spacing between lines on a piece of legal paper is 0.0087 m (8.7 mm). For each frame, find where the center of the penny is and estimate the number of lines from the top of the sheet that corresponds to that position. To get you started, the first two columns have been filled out for you. Look at Frames 1 and 2 to understand where these data come from, and then fill out the remaining columns.

<table>
<thead>
<tr>
<th>Frame</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>time (seconds)</td>
<td>0</td>
<td>1/23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position (lines from top)</td>
<td>0</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position (m)</td>
<td>0</td>
<td>0.01305</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Fill out the last four columns of this table to complete problem 1a

b) We can find the approximate speed of the penny by calculating the change in its position, and dividing by the change in time, $v = \Delta y/\Delta t$. For example, the change in position from Frame 1 to Frame 2 is $\Delta y = y_2 - y_1 = 0.01305 - 0m = 0.01305m$. The change in time from Frame 1 to Frame 2 is $\Delta t = t_2 - t_1 = 1/23 - 0sec = 0.04348sec$. These give $v = 0.01305/0.04348 = 0.30015$ m/s. Perform these calculations for your measurements from part 1a, and write your answers in the following table:

c) Plot your measurements and fit them with a straight line, using R. Here’s how:
Table 8: Fill out the last four columns of this table to complete problem 1b.

<table>
<thead>
<tr>
<th>Frame</th>
<th>1-2</th>
<th>2-3</th>
<th>3-4</th>
<th>4-5</th>
<th>5-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>time (sec)</td>
<td>1/23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Delta t ) (sec)</td>
<td>1/23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Delta y ) (m)</td>
<td>0.01305</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( v ) (m/s)</td>
<td>0.30015</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
t = \text{seq(from} = \frac{1}{23}, \text{to} = \frac{5}{23}, \text{by} = \frac{1}{23})
\]
\[
v = c(0.30015, 1, 1, 1, 1)
\]
\[
G <- \text{lm}(v \sim t)
\]
\[
\text{plot}(t, v)
\]
\[
\text{lines}(t, \text{fitted}(G))
\]

Note that you will have to modify the second line of this code, so that \( v \) stores your data for \( v \) from Table 2. To get the equation for the best-fit line, type in “G” at the prompt in R (or R-studio). If the linear fit is of the form \( y = mx + b \), then the coefficient called “intercept” is \( b \) and the other coefficient is the slope. What are these values?

d) You may recall from high school physics that the velocity of an object pulled by gravity is

\[
v(t) = v_0 + gt
\]

where \( v_0 \) is the initial speed and \( g \) is the gravitational constant. This equation says that your data, which estimate \( v(t) \), should be fit by a straight line with intercept \( v_0 \) and slope \( g \). Based on your answer to part 1c, what is \( g \)?

e) The penny’s acceleration is the change in speed with respect to time. What do you think that is?

f) How do you think you could make a more accurate measurement of the penny’s velocity, and of \( g \)?
Problem 2. Spread of a virus

The spread of a virus is modeled by

\[ I(t) = -t^2 + 6t - 4 \]

where \( I(t) \) is the number of people (in hundreds) infected with the virus and \( t \) is the number of weeks since the first case was observed.

a) Use R/R-studio to graph \( I(t) \).

b) What is a reasonable domain for \( I(t) \) (including units), i.e., what values of \( t \) makes sense as input for \( I(t) \)?

c) When does the number of infected individuals reach a maximum? What is the maximum number of cases?

d) Use the limit definition of the derivative to find the function which describes the rate of change of the number of infected individuals. Then find \( dI/dt \) at \( t = 2 \) and identify the units of this value. What do these units mean?

e) What is the rate of change in the number of cases at the maximum? How does this value compare with the phenomenon occurring at this time?

f) Give the sign (+ or −) of the rate of change up to the maximum and after the maximum.
Problem 1. MEASURING BICYCLE SPEED, I

It’s the weekend of the Great Davis Bicycle Speed Challenge (GDBSC) and, having heard of your reputation as a math whiz, the organizing committee wants your help. The main event in the GDBSC is a high speed bicycle race, where the fastest rider wins a home-made blueberry pie. Unlike most races, however, it is the rider who passes the finish line at the highest speed who wins, and not necessarily the rider who completes the course in the shortest time. The committee wants to know how to measure each rider’s speed as he or she crosses the finish line.

```
Start     Finish
___________ 1 mile __________
1 | Detector |
2 |
3 |
4 |
```

The committee has purchased two state-of-the-art measurement devices that can very accurately determine when a rider passes them. These can be arranged along the 1 mile-long track. In order to find the best arrangement of detectors, they are set up four different ways. One detector is fixed at the finish line, the other is set up: 1. at the starting line; 2. half-way down the track; 3. three quarters of the way down the track; and 4. seven eighths of the way down the track. The detectors then measure how long it takes the rider to complete the last mile, half mile, quarter mile or eighth of a mile of the race, respectively (see figure below). To test these arrangements, a practice rider races the course. The rider is specially trained to race the course exactly the same way each time. She starts from rest, and her speed increases throughout the course. Her times are listed in the table below.

<table>
<thead>
<tr>
<th>Detector Position</th>
<th>Starting Line (d)</th>
<th>Halfway (d)</th>
<th>Three Quarters (d)</th>
<th>Seven Eights (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured Time (t)</td>
<td>0.06138 hr</td>
<td>0.02144 hr</td>
<td>0.01023 hr</td>
<td>0.005027 hr</td>
</tr>
</tbody>
</table>

a) The rider’s apparent speed, $v_a$, is equal to the distance between detectors ($d$) divided by the measured time ($t$), so that $v_a = d/t$. Calculate the rider’s apparent speed for the four different detector arrangements.
b) Which measurement of the rider’s apparent speed, $v_a$, do you expect to be the most accurate (i.e. closest to the rider’s actual speed as she crosses the finish line)? Why?

c) Use R/R-studio to make a plot of the rider’s apparent speed, $v_a$, as a function of the spacing between detectors. Indicate your best guess of the rider’s actual speed on the plot.

d) You have made measurements of $v_a$ at several discrete values of $d$. Suppose you were able to determine a function that relates the two: $v_a(d)$. The rider’s actual speed, $v$, is a limit of this function, $v = \lim_{d \to c} v_a(d)$. Based on your answers to part b) and part c), what is the value of $c$ (i.e. what should $d$ approach in the limit)?

e) Suppose the racing committee let you place the detectors anywhere. How would you arrange them to best measure the speed of each rider as he or she crosses the finish line? Explain your answer (no credit will be awarded without an explanation).
Problem 2. MEASURING BICYCLE SPEED, II If you want to know how fast you’re going on your bike, you can buy a bicycle speedometer. Here’s how it works: As the bike moves, the wheel spins (clockwise, in the picture). A magnet, attached to the wheel, travels in a circle. It passes a detector, attached to the fork that does not spin. The detector outputs a signal (top), and we see a “blip” every time the magnet passes by the detector.

![Diagram of bicycle speedometer](image)

**a)** Suppose you measure the time between blips, $\Delta t_1$. Write an equation that would give the bike’s approximate speed. Your equation should use some or all of the variables $\Delta t_1$, $\Delta t_2$, $d$ and $r$ (defined in the picture).

**b)** Suppose you measure the time between blips, $\Delta t_2$. Write an equation that would give the bike’s approximate speed. Your equation should use some or all of the variables $\Delta t_1$, $\Delta t_2$, $d$ and $r$ (defined in the picture).

![Graph of bike speed](image)

**c)** This plot (above) shows the actual bike speed (solid line) and the two estimates of bike speed from part a) and part b). Which one is which? Explain your answer.

**Challenge question:** Explain why the hollow circles are always above the exact speed.
Problem 3. A DAMPED OSCILLATOR AND THE SANDWICH THEOREM

Perhaps you learned about a simple harmonic oscillator in physics class. This system consists of a small, heavy particle that is attached to a spring. The particle moves without friction. If you extend the spring and then let go, the particle slides back and forth.

This is an important example in physics class, but it is obviously an idealization. In the real world, there is friction. Here, you’ll examine a harmonic oscillator with friction – a “damped” oscillator. In order to give you an intuition for how one behaves, here’s how to build one:

Step 1: Take a sheet of paper and hold it in a “U” shape.
Step 2: Place a pen or pencil near the bottom of the “U”, but slightly off to one side.
Step 3: Let go, and watch the pen or pencil move back and forth.

Here’s a sketch of the set-up.

When you do this experiment, the horizontal position of the pen (x in the sketch on the right) is a function of time t, and approximately obeys an equation that looks like

\[ x(t) = e^{-t} \cos(\pi t) \]

a) Using R/R-studio, make a plot of \( x(t) \). Is that consistent with what you saw in the “experiment”? Explain.

b) What do you think happens after a long time? What do you think \( \lim_{t \to \infty} x(t) = \)? Explain.

c) The book shows you how to calculate this limit exactly (§3.4) using the “sandwich theorem.” The basic idea is as follows: the function \( \cos(\pi t) \) is never greater than 1. We can therefore find an upper bound for our function: \( x(t) = e^{-t} \cos(\pi t) \leq e^{-t} \) (where we’ve plugged in 1 for \( \cos(\pi t) \)).

The function \( \cos(\pi t) \) is never less than \(-1\). We can therefore find a lower bound for our function: \( x(t)e^{-t} \cos(\pi t) \geq -e^{-t} \) (where we’ve plugged in \(-1\) for \( \cos(\pi t) \)).

Combining these, we have

\[ -e^{-t} \leq x(t) \leq e^{-t} \]
Now, calculate the limit of the lower bound \((-e^{-t})\) and the upper bound \((e^{-t})\) as \(t\) goes to \(\infty\). We know that \(x(t)\) must lie between the two. So, what’s \(\lim_{t \to \infty} x(t)\)? Is that consistent with what you saw (part a) and what you expected to see (part b)? Explain.

d) There’s another way to figure out this limit. Suppose that you measure the \(x\)-position of the pen at \(t = 0, 2, 4, 6, \ldots\). This gives you a sequence of measurements, \(M_0, M_1, M_2, \ldots\). Write down the first five measurements. Then, write the “solution to the recursion,” that is, write an equation that gives you \(M_i\) for a given index \(i\). What is the limit of this sequence as \(i \to \infty\)? Is that consistent with your answers to parts a–c? Explain.
Week 5: Notes for the instructor

Purpose and summary

The purpose of this worksheet is to derive the product rule in an intuitive way.

There is only one problem on the worksheet. It goes through a geometric interpretation of the product rule, related to the derivation in the textbook on page 152. The idea is to look at the healing of a rectangular wound, that perhaps heals faster in one direction than the other. The students calculate the area symbolically and with numbers and then, at the end of the worksheet, derive the product rule.

Note that, since the wound is healing, the functions are decreasing and derivatives are negative. Keeping track of sign is a little tricky, so the TAs and LAs should make sure the students aren’t making sign errors.

Things to do in lecture

This worksheet corresponds to chapter 4.3 in Neuhauser, but can come before chapter 4.3 is discussed in lecture.

Materials

There is R/R-studio code (`Wound_Healing.R`) that shows a movie of the “wound” healing. This can be shown at the beginning of discussion. The movie has two aims: 1) to give the students a clear picture of the model system, and 2) to demonstrate some of the power of programming with R.
Week 5: Pre-discussion assignments
Problem 1. Wound healing

Suppose you are part of a research group tasked to determine the effectiveness of a drug that accelerates wound healing.

In the study, a surgeon gives a (brave) volunteer a rectangular wound on day 0, and the wound is photographed 10 times a day for the next three days. Since skin is not a uniform material, it heals faster in some directions than others. Knowing this, the surgeon orients the wound so that it heals faster along its height \( h(t) \) than its width \( w(t) \).

Since the drug might not have the same effect on healing along the wounds width and height, you decide to keep track of the wound’s area, \( A(t) = w(t)h(t) \).

Consider the following (idealized) pictures of the wound:

![Figure 5](image)

**Figure 5:** A picture of the wound at \( t = 1 \) day (left), and a picture of the wound at \( t + \Delta t = 1.1 \) day (right).

a) Fill out the following table, given that, when \( t = 1 \) day and \( t + \Delta t = 1.1 \) day, \( h(t) = 2 \) mm, \( h(t + \Delta t) = 1.9 \) mm, \( w(t) = 3 \) mm, and \( w(t + \Delta t) = 2.95 \) mm.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Symbol</th>
<th>In terms of ( h(t), w(t), \Delta t, ) and ( t )</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in height (mm)</td>
<td>( \Delta h )</td>
<td>( h(t + \Delta t) - h(t) )</td>
<td>-0.1 mm</td>
</tr>
<tr>
<td>Change in width (mm)</td>
<td>( \Delta w )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in time (days)</td>
<td>( \Delta t )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9: Fill out the remainder of this table to complete problem 1a.

b) The change in area, \( \Delta A = A(t + \Delta t) - A(t) \), is shown on the right of Fig. as the shaded region. It is made up of three rectangles. Give expressions for the areas of these three rectangles using only the variables \( h(t + \Delta t), w(t + \Delta t), \Delta h \) and \( \Delta w \), in
order to find an expression for $\Delta A$.

c) Check your answer to part b), by calculating the change in area using the given values of $h(t) = 2$ mm, $h(t + \Delta t) = 1.9$ mm, $w(t) = 3$ mm, and $w(t + \Delta t) = 2.95$ mm (recalling that $t = 1$ day and $t + \Delta t = 1.1$ day, ). Should the change in area be positive or negative?

d) Divide your expression for the change in area from part b) by $\Delta t$, which gives $\Delta A/\Delta t$, an approximation of the rate the area of the wound changes with time (why is this just an approximation?). Your answer should contain three terms, each one corresponding to the area of one of the shaded rectangles divided by $\Delta t$. Using the numbers you calculated in Table 10, evaluate each of the three terms. Is one much smaller than the others? Why do you think that is? Suppose you were to decrease $\Delta t$. What do you think would happen to each of the three terms?

e) Suppose we have a model for $h(t)$ and $w(t)$. That is, suppose, based on fits to the data, or an underlying physical law, we find that $h(t) = 3 - t$ and $w(t) = 3.5e^{-0.155t}$.

   i. Verify that these functions give the correct values for $h(1)$, $w(1)$, $h(1.1)$ and $w(1.1)$ used throughout this problem.

   ii. Recalling your expressions in Table 10, re-calculate $\Delta h$ and $\Delta w$ for $\Delta t = 0.01$, 0.001 and 0.0001 days. Plug these values into your answer for part d) (remembering to change $\Delta t$ when you divide) and plot the areas of the three terms as a function of $\log_{10}(\Delta t)$ using R/R-studio.

   iii. Does your plot support your answer to part d)? Discuss.

f) Now, take the limit as $\Delta t$ goes to zero in your answer to part d). Note that this makes

$$\lim_{\Delta t \to 0} \frac{\Delta A}{\Delta t} = \frac{dA}{dt}$$

the derivative of the wound’s area with respect to time. Each term in your answer to part d should also contain the derivative of a function with respect to time.

g) Based on your work above, what is

$$\frac{dA}{dt} = \frac{d}{dt} \left( w(t)h(t) \right) =$$
Discussion Questions (optional): Suppose that you are investigating two drugs that facilitate wound healing. You find the following models describe the width and height of the wounds:

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Equation for $w(t)$</th>
<th>Equation for $h(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>$w(t) = 3.5e^{-t0.16}$</td>
<td>$h(t) = 3 - t$</td>
</tr>
<tr>
<td>Drug 1</td>
<td>$w(t) = 3.5e^{-0.75t}$</td>
<td>$h(t) = 3 - 2t$</td>
</tr>
<tr>
<td>Drug 2</td>
<td>$w(t) = 3.5e^{-0.3t}$</td>
<td>$h(t) = 3 - 1.25t$</td>
</tr>
</tbody>
</table>

Table 10: Models describing wound healing in the presence and absence of drug treatment.

Since large wounds are dangerous, you are most interested in developing a drug that maximizes the initial rate that the wound closes. Which drug (if any) works the best? Notice that the initial wound had a width of $w(0) = 3.5$ mm and a height of $h(0) = 3$ mm. Does the best drug depend on the shape of the wound? Explain.

Note: It may be useful to know that \( \frac{d}{dx} (e^{ax}) = ae^{ax} \).
Problems for week 5: product rule, chain rule

1. Enzyme sensitivity

As you probably know, an enzyme is a protein that catalyzes a chemical reaction. Here, we consider a reaction where substrate, $S$, turns into product, $P$. The rate of product formation is called $v$.

The Hill model relates the rate of product formation in the presence of an enzyme ($v$) to the amount of substrate ($S$), according to the equation

$$v = V_{\text{max}} \frac{S^a}{K^a + S^a}$$

(6)

where $V_{\text{max}}$, $K$ and $a$ are positive constants. Notice that if $a = 1$, then this equation becomes the Michaelis-Menten equation, which we’ve seen a few times.

The purpose of this problem is to help you understand the parameter $a$, called the Hill coefficient. As we are less interested in the other parameters, we’ll set them equal to 1 (i.e. $V_{\text{max}} = 1$, $K = 1$).

Now, let’s consider three different enzymes, one with a Hill coefficient of 1

$$v_1 = \frac{S}{1 + S}$$

one with a Hill coefficient of 2

$$v_2 = \frac{S^2}{1 + S^2}$$

and one with a Hill coefficient of 4

$$v_4 = \frac{S^4}{1 + S^4}$$

Notice that, when $S = 1$, the three different enzymes have the same reaction rate $v_1 = v_2 = v_4 = 1/2$.

a) Calculate $\frac{dv_1}{dS}$, when $S = 1$.

b) Calculate $\frac{dv_2}{dS}$, when $S = 1$.

c) Calculate $\frac{dv_4}{dS}$, when $S = 1$.

d) Which enzyme is most sensitive to substrate concentration when $S = 1$? Explain.

(The sensitivity of an enzyme tells you the change in reaction rate that occurs when substrate concentration is changed).

e) Enzymes that control important biochemical reactions tend to have larger Hill coefficients. Why do you think that is?
Purpose and summary

The purpose of this worksheet is to look at an interesting application of the chain rule and quotient rule.

There are two problems on the worksheet. The second one is quite short, and can be eliminated, should time be a concern.

The first problem applies the chain rule to a simplified model of baseball pitching. The basic idea is for the students to see that throwing involves multiple “links” that rotate with respect to each other (i.e. the pitcher’s torso rotates with respect to his legs, his arm rotates with respect to his torso, etc.). The multi-link nature of throwing is important to maximize velocity. To understand this, two over-simplified models of throwing are introduced, one that has a single link and another that has two links.

The chain rule comes into this discussion when one tries to calculate the ball’s horizontal velocity given the angle and angular velocity of the link(s). In particular, since horizontal position is given by $x = \ell \sin(\theta)$ or $x = \ell_1 \sin(\theta) + \ell_2 \sin(\phi + \theta)$ for the two models, the horizontal velocity is given by $v_x = \ell \cos(\theta) \frac{d\theta}{dt}$ or $v_x = \ell_1 \cos(\theta) \frac{d\theta}{dt} + \ell_2 \cos(\phi + \theta) \left( \frac{d\theta}{dt} + \frac{d\phi}{dt} \right)$, respectively. Note that calculating the horizontal velocity for the second model might be a challenge for the students. TAs and LAs should be prepared to help with this step.

The velocity is maximized when the arm is straight ($\theta = 0$ or $\theta = \phi = 0$) and, since $\ell = \ell_1 + \ell_2$, the maximal horizontal velocities are $v_x = \ell \frac{d\theta}{dt}$ or $v_x = \ell_1 \frac{d\theta}{dt} + \ell_2 \frac{d\phi}{dt}$, for the one and two link systems, respectively. This result shows that the two link system gives a greater horizontal velocity.

The second problem is an application of the quotient rule (or the chain rule/product rule). The twist is that the students will need to set up the problem. In particular, they are asked to calculate the time rate of change of population density, given an area and its time rate of change, and a population and its time rate of change.

Things to do in lecture

This worksheet corresponds to chapter 4.4 in Neuhauser. The students will need to know the chain rule to complete problem 1, and the quotient rule (or the product/chain rule) to complete problem 2.

Materials

There are no materials for this week.
| **Week 6: Pre-discussion assignments** |
Problem 1. Optimal pitching

Let’s think about baseball. When a pitcher throws a ball, he steps forward, rotates his torso, rotates his shoulder with respect to his torso, rotates his forearm with respect to his upper arm, snaps his wrist and extends his fingers. The muscles controlling each of these body parts contract, and the timing is such that muscles controlling proximal (near the trunk) segments are activated before distal (near the finger tips) ones.

The multi-segment nature of this motion is extremely important for maximal velocity. In fact, the primary difference between a change-up and a fast-ball is that the change-up is gripped in the palm. This reduces the ability of the fingers to contract, which removes one segment from the motion, and ball velocity is about ten miles per hour slower than a fast-ball, where the fingers contribute to ball velocity.

In this problem, we’ll try to understand why multiple links help you throw faster.

Here are two simplified models of throwing:

In the model on the left, a single link pivots about a hinge. Suppose that we want to maximize the ball’s velocity in the $x$ direction (note the axes in the lower-left corner of the figure). Let’s write an expression for the ball’s velocity in the $x$ direction as a function of $\ell$, the segment length, $\theta$ the angle of the “arm,” and $\frac{d\theta}{dt}$ the rate that $\theta$ changes with time, called the angular velocity.

Here’s how we’ll do it:

a) Write an expression for the ball’s $x$-position in terms of $\ell$ and $\theta$.

b) Differentiate your expression from a) with respect to time. (Recall that $\theta(t)$ is a
function of time, and recall that \( \frac{d\theta}{dt} \) is the angular velocity.)

c) For a fixed value of \( \frac{d\theta}{dt} \), find the value of \( \theta \) that maximizes your expression for \( \frac{dx}{dt} \).

To do this you can graph your expression using R/R-studio or any other method. How is the “arm” positioned in order to get maximal ball velocity in the \( x \) direction? Does that make sense?

d) Suppose that the “arm” has length 1m, and the angular velocity \( \frac{d\theta}{dt} = 10 \) radians/second. What’s the maximum speed the ball can go? Write your answer in m/s, and also use the conversion factor 1m/s = 3.6mph to convert your answer to miles per hour (mph). Recall that radians are defined from the relationship between the arc length, \( s \), of a “wedge” of a circle with angle \( \theta \) and radius \( r \), given by \( s = r\theta \).

Therefore when the units of radians are used in the expression for angular velocity, the units of \( \frac{d\theta}{dt} \) are \( \frac{1}{\text{time}} \).

In the model in the right figure, one link pivots about a hinge and a second link pivots about the first. Again, suppose that we want to maximize the ball’s velocity in the \( x \) direction (note the axes in the lower-left corner of the figure). Let’s write an expression for the ball’s velocity in the \( x \) direction as a function of \( \ell_1 \) and \( \ell_2 \), the segment lengths, \( \theta \) the angle of the “upper arm,” its angular velocity \( \frac{d\theta}{dt} \) and \( \phi \) the angle of the “forearm” with respect to the upper arm, and its angular velocity \( \frac{d\phi}{dt} \).

e) repeat steps a–c for the model at the right. Your final answer will include the variables, \( \ell_1, \theta, \ell_2, \phi, \frac{d\theta}{dt} \) and \( \frac{d\phi}{dt} \).

f) Now suppose that the “arm” has same length as before (1m), and the joint occurs in the middle, giving \( \ell_1 = 0.5 \) m, and \( \ell_2 = 0.5 \) m. Again, suppose that the angular velocity \( \frac{d\theta}{dt} = 10 \) radians/second, but now the distal link also has angular velocity \( \frac{d\phi}{dt} = 10 \) radians/second. What’s the maximum speed the ball can go? Write your answer in m/s, and also use the conversion factor 1m/s = 3.6mph to convert your answer to miles per hour (mph).

g) Suppose that the model on the left (parts a-d) represents a change-up, while the model on the right (parts e and f) represents a fast-ball. Does that make sense? How does that second link contribute to ball speed?
Problem 2. Population Density

Shenzhen, which is a major city in southern China’s Guangdong Province, is one of the fastest growing cities in the world. In 1996, the area of Shenzhen was 600 km$^2$ and it was expanding at a rate of 40 km$^2$/year. The population of Shenzhen in 1996 was 5,000,000 people, and it was growing at a rate of 400,000 people/year.

a) Determine the population density of Shenzhen in 1996.

b) Use the data above to determine how fast the population density of Shenzhen was growing in 1996.

(Include the appropriate units).
Problems for week 6: chain rule, quotient rule
Purpose and summary

This worksheet shows some applications of linearization. There are two problems, and problem 1 is probably the best to eliminate, if time is a concern.

The purpose of this worksheet is to give the students an idea of why linearization is a critically important concept. We believe that this needs to be emphasized, since they will be seeing linearization again in 17B and in 17C. If they haven’t seen why it’s so important, then they are less likely to remember it.

Problem 1 deals with error propagation. This is a (more or less) standard problem that relates pretty closely to the book’s discussion of error propagation in chapter 4.8. One difference is that the student is encouraged to think about units.

Problem 2 shows a different application of linearization. They are introduced to the logistic model of population growth (which we return to in worksheets 9 and 10). They calculate the two fixed points and observe, using R, that the model settles down to the non-zero fixed point $K$, the carrying capacity.

In the next part, a (small) term for hunting pressure is added to the model. With hunting, it becomes impossible to find the fixed point. However, by linearizing about the fixed point for the model without hunting, $K$, the new fixed point can be estimated. After the students estimate this new fixed point, they check their answer with R.

Things to do in lecture

This worksheet corresponds to chapter 4.8. It would be possible to do the worksheet before linearization is covered in lecture, but it would probably best to eliminate one of the problems.

Materials

There are no materials for this discussion sheet.
Week 7: Pre-discussion assignments
Problem 1. Metabolic Rate for a California Condor

In many biological situations, we directly measure one quantity, but we really want to know a function of that measured quantity. For example, we might measure the partial pressure of oxygen in the blood ($p$) but want to know the saturation of hemoglobin, $S$, which is given by the hemoglobin saturation equation ($S = f(p) = 1/(1 + e^{-ap})$, where $a$ is a positive constant that depends on temperature, pH and other properties of the blood).

When using real data to predict the output of a function, measurement error for the independent variable (say, $x$) leads to an error in the estimate of the dependent variable (say, $y$). This phenomenon, known as error propagation, can be estimated using linear approximation.

a) We can approximate a function near a real number $x = a$ using the linear approximation of $f$ around $a$:

$$f(x) \approx f(a) + f'(a)(x - a)$$

i) Use the linear approximation of $f(x)$ at $x = a$ to fill in the following lines:

$$\Delta y = f(a + \Delta x) - f(a)$$

$$\approx \ldots - f(a)$$

$$= \ldots$$

Note that in this particular context, the term $f'(a)$ in the final expression is referred to as the sensitivity of $y$ to $x$ at $x = a$.

ii) Conceptually, why do you think we refer this term as sensitivity? What would a larger (or smaller) sensitivity value change about our linear approximation at a particular $x = a$?

b) To examine error propagation in more detail, let’s consider a model for the metabolic rate of animals. The following curve models how the metabolic rate $R$ (in kilocalories/day) depends on body mass $M$ (in kilograms):
\[ R = e^{4.2}M^{0.75} \]

i) Use this equation to predict the metabolic rate of a California Condor weighing 10kg.

ii) Naturally, not every California Condor weighs exactly 10kg, some weigh more and some weigh less. So, what is the sensitivity of our model to this measurement? Using your equation from part a.i, suitably modified so that body mass \( M \) is the independent variable, metabolic rate \( R \) is the dependent variable and the functional relationship between the two is \( R = e^{4.2}M^{0.75} \). Select a small error \( \Delta M \) and see how it propagates to an error \( \Delta R \) in our estimate of \( R \).

iii) What are the units of this propagated error, \( \Delta R \)?

**Error vs Percent Error**

c) Note that in parts a and b above, we were discussing \( \Delta R \), which contains units. It’s often more convenient to consider relative error, which does not contain units. That is, suppose that we know that most adult Condors weigh within 10% of 10kg. Then, we would know that \( \Delta M/M = 0.1 \).

i) Given a 10% error in the estimate of the condor’s weight, what (approximate) percent error will we have in the metabolic rate? (Hint: if percent error in weight is \( 100 \cdot \Delta M/M \), what do you think the percent error in metabolic rate is?)
Problem 2. Estimating a Fixed Point in a Population Model

Recall our discussion of recursions from a few weeks back. We can model, say, the population of animals at year $i$, $N_i$, using a simple equation like

$$N_{i+1} = 1.04N_i$$

given $N_0 = 25$. This works well for small populations that are not limited by resources – e.g. you saw that it works well for Whooping Cranes, which had nearly been driven to extinction in the 1940s.

However, these exponential growth models quickly become inaccurate, because exponential growth is just too fast. E.g., in the above example, the population doubles every 17.67 years, so that after 176.7 years, there would be $25 \cdot 2^{10} = 25,600$ individuals, and after 706.92 years ($17.67 \cdot 40$), there would be $25 \cdot 2^{40} = 27.5$ trillion individuals. Clearly, we can’t sustain a population of 27.5 trillion cranes.

A more realistic population model is the following, which includes resource-limited growth

$$N_{i+1} = N_i + aN_i(K - N_i)$$

This model has two fixed points and, after a long time, the population settles down to one of them (i.e. $\lim_{i \to \infty} N_i$ exists, unlike in the simple exponential growth model).

a) Find the two fixed points

b) Using R/R-studio, and the following code as a template, find $N_i$ for $i = 0, 1, 2, \ldots, 49, 50$, and plot $N_i$ vs. $i$. For this part, assume that $a = 0.04$, $K = 15$ and $N_0 = 2$. Has the population settled down to one of the fixed points? Which one?

This code simulates the simple exponential growth model ($N_{i+1} = 1.04N_i$, with $N_0 = 25$) for $i = 0, 1, 2, \ldots, 49, 50$ and then plots the results. You will have to modify this code in order to complete parts b and e.

```
N=25
for(i in 1:50){
N[i+1]=1.04*N[i]
}
plot(seq(from=0,to=50,by=1),N)
```
The previous model describes resource-limited growth, but suppose we want to include an effect like hunting in our model. Hunters remove individuals from the population if the population is sufficiently large. If the population is too small, then hunters seldom encounter their quarry, and hunting has no effect.

\[ N_{i+1} = N_i + aN_i(K - N_i) - h(1 - e^{-N_i/K}) \]

Notice that you can no longer solve for the fixed points. However, if hunting is sustainable, then the last term should be small compared to the others. Thus, we might expect that the fixed points are near the fixed points you found in part a.

Here’s where linearization can be of use. We can linearize the equation about the fixed point \( N_i = K \), and then solve this linearized equation to approximate the fixed point. We’ll do this in two steps:

**c) Step 1:** Linearize the function

\[ f(N_i) = N_i + aN_i(K - N_i) - h(1 - e^{-N_i/K}) \]

near \( N_i = K \).

(Recall that we can approximate a function near a real number \( N_i = a \) using the linear approximation of \( f \) around \( a \): \( f(N_i) \approx f(a) + f'(a)(N_i - a) \)).

**d) Step 2:** Now, the recursion is \( N_{i+1} = \) the linear function of \( N_i \) from part c.

Find the fixed point of this new recursion to approximation the fixed point of the original recursion. The algebra is messy. To simplify your work, assume that \( a = 0.04 \), \( K = 15 \), \( h = 0.1 \) and \( N_0 = 2 \).

Note: Show your work! To check your answer, the general formula is

\[ N = K - \frac{h(1 - e^{-1})}{aK + h/(eK)} \]

**e) How well does this estimate work?** Using R/R-studio, evaluate the exact, non-
linear recursion to get $N_i$ for $i = 0, 1, 2, \ldots, 49, 50$, and plot $N_i$ vs. $i$. For this part, assume that $a = 0.04$, $K = 15$, $h = 0.1$ and $N_0 = 2$. What is $N_{50}$? How does it compare to the fixed point you calculated in part c)?

Discussion questions: Suppose you’re managing a population of animals, and you have to set a limit on hunting. According to the general equation for the fixed point,

$$N = K - \frac{h(1 - e^{-1})}{aK + h/(eK)}$$

how much hunting do you think is sustainable (i.e. how big a value of $h$ would you allow)? Using R/R-studio, simulate the population for 50 years and see if it’s sustainable. Was your choice good? Explore a few values and see if anything interesting happens (HINT: our linear approximation assumes that hunting pressure is small).
Problems for week 7: linearization

Problems 1 and 2 share a common theme. In them, you will see how physicists, chemists, biologists and engineers typically use calculus. Problem 1 describes a method to solve differential equations (you’ll learn more about these in Math 17B and C). This method is based on an approximation of the derivative. Problem 2 describes how to find where a function is equal to zero. This method is based on linearization.

To understand why these methods are important, let us consider two different models, the first from biology and the second from physics. In problems 1 and 2, you will choose one of these to examine in more detail.

Model 1: Extinction of an endangered species. Suppose we are interested in the number of individuals of a particular endangered species, \( N(t) \). Further, suppose that the population is governed by the following differential equation

\[
\frac{dN}{dt} = -aN + \frac{bN^2}{1 + N^2} - h
\]

As you can see, on the left hand side is the change in the population with time.

On the right, the first term, \( aN \), is the death rate. That is, each year, some proportion of the population (represented by the constant \( a \)) dies from natural causes, including predation, sickness, old age, etc.

The next term, \( bN^2/(1 + N^2) \), is the birth rate. This is a somewhat complex birth rate, but it’s what you might expect if, say, there are a finite number of nesting sites for our endangered animal. Then, if there are a lot of animals, all of the nest sites are occupied and \( b \) animals are born each year. Alternatively, if there are only a few animals, the birth rate depends on the likelihood of two of the animals finding each other, which is proportional to \( N^2 \).

The final term, \( h \), is the number of animals harvested each year by human hunters. As you can see, each term associated with death has a minus sign (since these terms cause the number of animals to decrease), while the term associated with birth has a plus sign (since this term causes the number of animals to increase).

Model 2: Trajectory of a projectile. Suppose that you throw a ball straight up. The ball has initial velocity \( v_0 \). As the ball flies through the air, it feels a drag force. That drag force opposes the motion, with a magnitude proportional to the square of the velocity. It turns out that the equations that describe the height of the ball as a function of time are the following:

\[
\frac{dy}{dt} = \begin{cases} \frac{\sqrt{g}}{c} \tan \left(-c\sqrt{gt} + \tan^{-1} \left( \frac{cN}{\sqrt{g}} \right) \right) & : \frac{dv}{dt} \geq 0 \\ \frac{\sqrt{g}}{c} \tanh \left(-c\sqrt{g}(t - t_s) \right) & : \frac{dv}{dt} < 0 \end{cases}
\]

where \( g \) is the gravitational constant, \( c \) a constant that determines the strength of the drag force, and \( t_s \) is the time that the ball switches from going up (\( dy/dt > 0 \)) to going down (\( dy/dt < 0 \)).
1. **Forward Euler.**
Both Model 1 and Model 2 are described by differential equations. How do we deal with them? Here, we’ll learn a simple method to approximate their solution.

To see how it works, let’s consider a differential equation that’s much simpler than either Model 1 or Model 2. Here it is:

\[
\frac{dx}{dt} = -kx
\]  
where \( k \) is a positive constant.

**a)** Eq. 12 has a solution. By directly evaluating the derivative, \( dx/dt \), show that

\[
x(t) = e^{-kt}
\]
satisfies it.

**b)** We can approximate this solution with a sequence of points, \( X_i \). Here’s how. First, we re-write Eq. 12 using the definition of the derivative

\[
\lim_{\Delta x \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} = -kx(t)
\]
then, we drop the limit, which will make the equation only approximately true (although the approximation can be very good if we pick a small \( \Delta t \))

\[
\frac{x(t + \Delta t) - x(t)}{\Delta t} \approx -kx(t)
\]
now, doing a little algebra, we can rearrange the equation

\[
x(t + \Delta t) \approx -kx(t)\Delta t + x(t)
\]

Remember writing sequences from functions evaluated at evenly spaced intervals? Suppose that we do that with \( x(t) \). That is, suppose \( X_0 = x(0) \) and \( X_1 = x(\Delta t) \) and \( X_2 = x(2\Delta t) \) and so on. Using Eq. 8 write a recursion that defines the sequence \( X_i \). (Note: a recursion requires an initial condition. Here, you may assume that \( X_0 = 1 \)).

**c)** The sequence that you defined in part b) is an approximation of the exact solution (which you found in part a). How good is it? Let’s find out. Suppose \( k = 1 \). Use R/R-studio (and maybe a “for” loop) to evaluate the first 20 terms in the sequence \( (X_0, X_1, ..., X_{19}) \) with \( \Delta t = 0.01 \). Plot these along with the exact solution you found in part a.

**d)** Choose either Model 1 or Model 2. Describe how you would generate a sequence that approximates the exact solution.
2. **Newton’s Method.**

For the biological system described in Model 1, we are interested in when \(N(t) = 0\). That is, we want to know when the endangered species will go extinct.

For the physical system described in Model 2, we want to know when the sign of the upward velocity switches from being positive \((dy/dt > 0)\) to being negative \((dy/dt < 0)\). We called this time \(t_s\) and, at that time, \(dy/dt = 0\).

Thus, in both Model 1 and Model 2, we want to find when something goes to 0. We can use linearization to find when a function goes to 0. The technique is called Newton’s Method.

As in problem 9, it will be useful to consider a very simple example to see how the method works. Suppose we want to find a root of the function \(y(x)\) — that is, we want to find where \(y(x) = 0\).

Here’s the function we’ll consider:

\[
y(x) = -x^3 + x^2 + 1
\]

The idea of Newton’s method is that if we can guess a number near that root, then the linear approximation will be very good. We can use this linear approximation to predict the value of the root.

Look at the following plot. It shows the function, and you can see it passes through 0 near \(x = 1.5\).

![Figure 6](image)

**Figure 6:**

a) Find the linear approximation of \(y(x)\) about \(x = 1.5\).

b) Use this linear approximation to predict the value of \(x\) where \(y(x) = 0\). (Hint: your linear approximation will look something like \(a(x - 1.5) + b\). To predict the value of the root, set this equal to 0, and then solve for \(x\).)

c) Now, repeat the process, starting from your answer to part b. (That is, linearize about \(x = \) [your answer to part b]. Then, solve for the new predicted value of the root).
d) The actual value of the root is \( x = 1.46557123 \ldots \). Is your answer to part b) or part c) closer to the actual value? Why do you think that is? If you repeated the process, do you think that estimate would be better or worse? Discuss.

e) Choose either Model 1 or Model 2. Describe how you would use Newton’s Method and the Forward Euler Method to find when the endangered species goes extinct (Model 1) or the height of the ball as a function of time (Model 2). I wrote R/R-studio programs that simulate both models. You can run the programs and try to see where I use Forward Euler and Newton’s methods.
3. DAMPED OSCILLATOR, REVISITED In week 4, you may have done a problem about a damped oscillator. If you didn’t do it last time, or if you want to see it again, here’s how to build one:

Step 1: Take a sheet of paper and hold it in a “U” shape.
Step 2: Place a pen or pencil near the bottom of the “U”, but slightly off to one side.
Step 3: Let go, and watch the pen or pencil move back and forth.

Here’s a sketch of the set-up.

When you do this experiment, the horizontal position of the pen ($x$ in the sketch on the right) is a function of time $t$, and approximately obeys an equation of the form

$$x(t) = x_0 e^{-at} \cos(bt)$$

where $x_0$ is the initial horizontal position of the pen, and $a$ and $b$ are positive constants. Here, we’ll think about why this happens and get a physical understanding of $a$ and $b$.

In physics class, you might have learned about a damped harmonic oscillator. This is a system where a particle, of mass $m$, is attached to a linear spring and a linear dashpot. A linear spring applies a force that is proportional to the position of the mass ($F_s = -kx$), while a linear dashpot applies a force that is proportional to the velocity of the mass ($F_d = -d\frac{dx}{dt}$). Then, Newton’s second law for the system is

$$F = ma$$

$$-kx - d\frac{dx}{dt} = m\frac{d^2x}{dt^2}$$

You will now show that, with suitable choices for $a$ and $b$, equation 9 is a solution to this differential equation, and that this differential equation describes the system you built when $x$ and $\frac{dx}{dt}$ are small. To do so:

a. Differentiate Eq. 9 with respect to time (i.e. calculate $\frac{dx}{dt}$).

b. Differentiate Eq. 9 twice with respect to time (i.e. calculate $\frac{d^2x}{dt^2}$).
c. Re-write your answer to part b to be in terms of $x$ and $\frac{dx}{dt}$. Your final answer should look like

\[
\frac{d^2x}{dt^2} = A \frac{dx}{dt} + Bx
\]

where $A$ and $B$ depend on the constants $a$ and $b$.

d. Your answer to part c should be of the same form of the original differential equation, Eq. 10. Find an expression for $a$ and $b$ in terms of the physical constants $k$, $d$ and $m$. Then, substitute these into the solution to the differential equation, Eq. 9.

e. In part d, you showed that a harmonic oscillator moves back and forth according to Eq. 9. But the physical system you built is not exactly a harmonic oscillator. Instead, it has a force that depends on position $F_s = f(x)$ and a frictional force that depends on velocity $F_d = g(\frac{dx}{dt})$. Given that the spring force and frictional force are both 0 when their independent variables are 0, briefly explain why Eq. 10 is a good approximation when $x$ and $\frac{dx}{dt}$ are small.

f. Given your answer to e, what are $k$ and $d$ in terms of of the spring force function $f(x)$ and the damping force function $g(x)$?
Purpose and summary

This worksheet deals with optimization and graphing. There are two problems, either one can be eliminated.

In the first problem, we return to the problem of the population of whooping cranes from worksheet 3. In that problem, we used R to find the best-fit line to the data on a semi-log plot. Here, the students redo that problem, but using optimization to find the line with the best-fit slope.

To make the problem simpler, only three data points are used. The line is assumed to pass through the first data point, so there is a single parameter with which to find the best-fit line: the slope, $m$. As a first step, the students determine what a reasonable measure of error ($E$) should be. Next, they determine how to write the error as a function of the slope of the fitting line, $E(m)$. Finally, to minimize this quantity, they set $\frac{dE}{dm} = 0$ and solve. To check their answer, they plot the best fit line along with the data.

In the second problem, a dose-response curve is provided. The students sketch the function. However, in order to sketch the function, they are asked to think about the function in terms of dosage (e.g., instead of asking whether the function has a horizontal asymptote, the student is asked “what happens if you take an extremely large dose”). The idea here is to get the students to think about the process of graphing as a series of logical steps, rather than a recipe to follow.

Things to do in lecture

Problem 1 in this worksheet corresponds to the following sections, 5.1 extrema and the mean value theorem, 5.3 extrema, and 5.4 optimization. It can be introduced early in chapter 5 (i.e. before optimization is covered), if the TA/LAs guide the students through part e.

Problem 2 in this worksheet corresponds to the section on graphing (chapter 5.3), but can be introduced early as a motivation for that section.

Materials

There are no materials for this discussion sheet.
Week 8: Pre-discussion assignments
Problem 1. Conservation of whooping cranes, revisited Recall Worksheet 3, where you looked at the number of wild Whooping Cranes each decade from 1940 to 2010. It comes from a paper by Butler, Harris and Strobel, published in the journal Biological Conservation in 2013.

Table 11: Data from Butler et al. 2013

<table>
<thead>
<tr>
<th>Year</th>
<th>time since start (t) in years</th>
<th>Number of Cranes (N)</th>
<th>log_{10}(N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940</td>
<td>0</td>
<td>26</td>
<td>1.41</td>
</tr>
<tr>
<td>1950</td>
<td>10</td>
<td>32</td>
<td>1.51</td>
</tr>
<tr>
<td>1960</td>
<td>20</td>
<td>37</td>
<td>1.57</td>
</tr>
<tr>
<td>1970</td>
<td>30</td>
<td>58</td>
<td>1.76</td>
</tr>
<tr>
<td>1980</td>
<td>40</td>
<td>79</td>
<td>1.90</td>
</tr>
<tr>
<td>1990</td>
<td>50</td>
<td>146</td>
<td>2.16</td>
</tr>
<tr>
<td>2000</td>
<td>60</td>
<td>180</td>
<td>2.26</td>
</tr>
<tr>
<td>2010</td>
<td>70</td>
<td>283</td>
<td>2.45</td>
</tr>
</tbody>
</table>

These data were linear on a semi-log plot; that is, when you plot time since start, $t$ (column 2) vs. $\log_{10}(N)$, (column 4), the plot is linear. The slope of the best-fit line tells you the population growth rate per year, but how do you find the best-fit line?
In Discussion #3 you used R/R-studio to find the best-fit line using the following commands, which give the figure above.

`t=seq(from = 0, to = 70, by = 10)
N=c(26,32,37,58,79,146,180,283)
F<-lm(log(N,base = 10)~t)
plot(t,log(N,base = 10))
lines(t,fitted(F))

The built-in function `lm` in R/R-studio will find this best-fit line; in this problem you will see how it works.

To simplify the problem, let’s fit the data with a straight line that passes through the first point. That is, let’s assume that the data are fit by a model of the form \( \log_{10}(N) = mt + 1.41 \). We want to pick \( m \) so that this line fits the data best.

To set up our notation, for the \( k^{th} \) time value \( t_k \), let the \( \log \) of the actual number of cranes be \( N_k \), and the \( \log \) of the predicted number of cranes be \( \hat{N}_k \), where our model is

\[
\hat{N}_k = mt_k + 1.41.
\]

Let’s start with a simpler table, one that has only three entries and only contains the relevant data,

<table>
<thead>
<tr>
<th>( t_k )</th>
<th>( N_k )</th>
<th>( \hat{N}_k = mt_k + 1.41 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1=0 )</td>
<td>1.41</td>
<td></td>
</tr>
<tr>
<td>( t_2=40 )</td>
<td>1.90</td>
<td></td>
</tr>
<tr>
<td>( t_3=70 )</td>
<td>2.45</td>
<td></td>
</tr>
</tbody>
</table>

a) Fill in the last column of the table (except for your entry in the first row, each of your answers will contain the variable \( m \)).

**Defining an appropriate measure of “total error”**

We define the *error* of the model for the \( k^{th} \) input value to be the difference between the actual output and the predicted output

\[
e_k = N_k - \hat{N}_k = N_k - (mt_k + 1.41),
\]

e.g., for \( k = 3 \), the actual output value is \( \hat{N}_3 = 2.45 \), the output predicted by the model is \( \hat{N}_3 = m70 + 1.41 \), and the error is \( e_3 = 2.45 - (m70 + 1.41) \).

To find the best fit of our model to the data we want to minimize the *total error*. However, what is a appropriate measure of total error?
One might think that that total error should be defined as the sum of all of the errors from the individual data points

$$e_{1,\text{tot}} = e_1 + e_2 + e_3.$$ 

However, there are serious problems with this definition of total error.

(b) Suppose we propose the line $N_k = 10.0t_k + 1.41$ as our model for the three data points, i.e. $m = 10.0$. Compute the total error for this model, using the sum of the individual errors $e_{1,\text{tot}}$ to define total error. What’s wrong with this definition of total error? (Things to think about: is the fit good? why not? you want to minimize the error, is the error small?)

There are at least two ways to remedy the situation. One would be to use sum of the absolute value of the individual errors $e_k$; another would be to use the sum of squares of the individual errors. For mathematical reasons, the squaring approach is usually used.\footnote{This is related to the fact that we define the distance from two points using the sum of squares, i.e. $d((x_1, x_2), (y_1, y_2)) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$, this comes from the Pythagorean theorem.} Thus, we can define \textit{total squared error} as

$$e_{2,\text{tot}} = e_1^2 + e_2^2 + e_3^2.$$ 

(c) For the given data and the suggested line $N_k = 10.0t_k + 1.41$, compute the total error, using the sum of the squares of the individual errors to define total error. Why is this definition of total error better than our previous one?
Minimizing the total error

(d) Now, we want to find the line, \( N_k = mt_k + 1.41 \), that minimizes the error for our three data points

\[
(t_1, N_1) = (0, 1.41), \quad (t_2, N_2) = (40, 1.90), \quad (t_3, N_3) = (70, 2.45),
\]

Convince yourself that the sum of the squared error, \( e_{2,\text{tot}} \) is

\[
e_{2,\text{tot}} = e_1^2 + e_2^2 + e_3^2 = (1.41 - (0m + 1.41))^2 + (1.90 - (40m + 1.41))^2 + (2.45 - (70m + 1.41))^2 = 0 + (0.49 - 40m)^2 + (1.04 - 70m)^2 = 1.3217 - 184.8m + 6500m^2 = E(m)
\]

(e) For what value of \( m \) is the the sum of the squared error \( e_{2,\text{tot}} = E(m) \) minimized, (i.e., what \( m \) gives the best fit to the data)?

(f) Use R/R-studio to plot the data along with the best fit line. Is it a reasonable fit? Compare this best-fit line to the one that R/R-studio gives using the command `lm`. Are they different? If so, why do you think that is?
Problem 2. Dose-Response of Multi-Vitamins

Multi-vitamins typically have dose-response curves of the following form

\[ R = f(x) = \frac{ax}{k^2 + x^2}, \quad x \geq 0, \]

where \( x \) is a measure of the daily dose, \( R \) is a measure of the health benefits, and \( a \) and \( k \) are positive constants.

(a) According to this model what is the health benefit to taking no vitamins? How does this change if you take a very small amount?

(b) What happens if you take an extremely large dose?

(c) What is the range of dosage where taking more will increase the benefit, and what is the range of dosage where taking more will decrease the benefit?

(d) Is there ever a negative benefit to taking the vitamins, according to this model?
(e) Is there a dosage that maximizes or minimizes the health benefits?

(f) Graph the function and interpret your graph in terms of the health benefits of the multi-vitamins.
Problems for week 8: extrema, graphing, optimization

1. Consider the following chemical reaction

\[ A \xrightarrow{k_1} B \xrightarrow{k_2} C \]

If the two reaction rates \((k_1 \text{ and } k_2)\) are not equal, the concentration of \(B\) as a function of time is

\[ B(t) = \frac{k_1}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t}) \]

Find both the maximum value of \(B\) and the time it reaches this maximum.

2. Suppose you throw a stick in the water for your dog. Your dog will have to run across the land and then swim through the water to reach the stick. On land, your dog runs 12 m/s. In the water, your dog swims 1 m/s. What path should your dog take in order to minimize the time it takes her to reach the stick?

Assume that the stick lands a distance \(y\) in front of you and \(x\) to your right. Assume that the dog travels in a straight line to the shore, and then another straight line (but perhaps in a different direction) in the water to reach the stick. Assume that the shore is a straight line, parallel to the \(x\)-axis, and starts at \(y/2\) (see the diagram, below).

To describe the optimal path, find the value of \(L\) that minimizes the time it takes your dog to reach the stick.\(^2\)

NOTE: start by setting up the problem, so that you get an equation like \(f(L) = 0\). You’ll get an equation that you can’t solve (called a transcendental equation), but you can plot it in R/R-studio and (approximately determine the value of \(L\) at which \(f(L) = 0\).

\(^2\)It may interest you to know that Professor Tim Pennings at Hope College, in Michigan, showed that his dog Elvis approximately chose this optimal path to retrieve a ball thrown in the water – you can look it up on the internet to learn more.
3. **Determining Optimal Dosage.**

In toxicology, it is often said that the difference between a cure and a poison is simply dosage\(^3\). The idea is that a medicine, if taken in too large a dose can be fatal. Obviously however, if untreated, some diseases are fatal. Particular drugs, if taken in an appropriate amount, lead to a higher probability of survival. But, how does one determine the correct dose? The answer depends on the severity of the patients illness, and the toxicity of the medicine.

**a.** Suppose you are trying to determine the correct dose of a drug. You know that the percentage of patients that die \((P_T)\) at a given dosage \((D)\) due to drug toxicity is

\[ P_T = p_T \cdot 100\% = \frac{1}{4} D \cdot 100\% \]

for dosages between \(0 \leq D \leq 4\) (let’s not worry about units). That is, given no drug \((D = 0)\) no patients die from the toxicity of the drug; however, given a very large dose of the drug \((D = 4)\) all of the patients die.

Naturally, a healthy person would not take the drug, since even a low dose carries a risk of death. However, this drug can cure a potentially fatal disease. The drug works by helping the patient’s immune system kill the pathogen that causes the disease. Suppose that the pathogen is killed \((P_K)\) at a given dosage \((D)\) according to the equation

\[ P_K = p_k \cdot 100\% = (0.1 + \frac{1}{5} D) \cdot 100\% \]

for dosages between \(0 \leq D \leq 4\). That is, given no drug \((D = 0)\) the patient’s immune system kills off the pathogen 10% of the time; however, given a very large dose of the drug \((D = 4)\) the patient’s immune system kills off the pathogen 90% of the time. The percentage of patients that survive \((P_S)\) at a given drug dosage \((D)\) is the probability that the pathogen is killed times the probability that the drug does not kill them:

\[ P_S = p_K(D) \cdot (1 - p_T(D)) \cdot 100\% = (0.1 + \frac{1}{5} D)(1 - \frac{1}{4} D)100\% \]

**USE CALCULUS** to determine:

i. the dosage that maximizes the percentage of patients that survive (i.e. find \(D\) to maximize \(P_S\))

ii. the percentage of patients that survive at this dosage (i.e. find \(P_S\) at the value of \(D\) from part i.).

**b)** It is very difficult to measure \(P_T\) and \(P_K\) in humans. Typically, after a drug is developed, it goes through Phase I clinical trials where it is tested on animals. At this stage, the functions \(P_T\) and \(P_K\) are measured. They are rarely as simple as for part a. More realistically, one might find

\[ P_T = p_T \cdot 100\% = \frac{D}{5 + D} \cdot 100\% \quad P_K = p_k \cdot 100\% = \frac{D^2}{2^2 + D^2} \cdot 100\% \quad (11) \]

\(^3\)Apparently, the original quote is from Paracelsus, a 16th century German-Swiss thinker: “All things are poison, and nothing is without poison; only the dose permits something not to be poisonous.”
as in part a, the percentage of animals that survive ($P_S$) at a given drug dosage ($D$) is the probability that the pathogen is killed times the probability that the drug does not kill the animal:

$$P_S = p_K(D) \cdot (1 - p_T(D)) \cdot 100\% = \frac{D^2}{2^2 + D^2} \cdot \frac{5}{5 + D} \cdot 100\%$$

USE R/R-STUDIO to determine (to receive full credit, you must sketch the graph):

i. the dosage that maximizes the percentage of animals that survive, and
ii. the percentage of animals that survive at that dosage.

c) The dosage from Phase I clinical trials is only a guess. In Phase II and Phase III clinical trials, one tries to refine this guess by performing human trials. Suppose that the suggested dosage from Phase I clinical trials is $D*$.

Let's say you find that $p_K(D^*) = 0.5$ and $p_T(D^*) = 0.5$ for a group of patients. Suppose you then change the dosage by a small amount and evaluate these values for a different group of patients. In particular, you find $p_K(D^* + \Delta D) = 0.51$ and $p_T(D^* + \Delta D) = 0.505$ for $\Delta D = 0.01$. Use these values to find a linear approximation of $p_K(D)$ and $p_T(D)$. Then, using the equation $P_S = p_K(D) \cdot (1 - p_T(D)) \cdot 100\%$

i. find the dosage that you think will maximize the percentage of patients that survive and
ii. estimate the percentage of patients survive at this dosage.

**Hints for part c:**

1. Use the values of $p_K(D^* + \Delta D)$ and $p_K(D^*)$ to estimate the derivative of $p_K$ with respect to $D$. Do the same for $p_T$. It might help to recall that $\frac{df}{dx} \approx \frac{f(x+\Delta x) - f(x)}{\Delta x}$
2. Recall that the formula for linearizing a function near $x = a$ is $f(x) \approx f(a) + \frac{df}{dx} \bigg|_{x=a} (x-a)$
3. Look at your solution to part a
4. Your calculation of the dosage will include $D^*$, your estimate of the percentage of patients that survive will not.
5) An enzyme that exhibits negative cooperativity has a reaction rate $v$ as a function of substrate concentration $s$ of the form

$$v = V_{\text{max}} \frac{s}{K^a + s^a}$$

where $V_{\text{max}} > 0$, $K > 0$ and $a > 1$ are constants. Find the substrate concentration, $s_{\text{max}}$, at which reaction rate is a maximum.

Note: you should assume the $s \geq 0$, and to get full credit, make a brief argument for why the point you found is a maximum.
Purpose and summary

This worksheet deals with stability of fixed points of difference equations. This is the first of two worksheets on this topic. The idea is to emphasize how linearization is used to determine stability, which the students will see again in 17B and 17C. Note that it is possible to do this worksheet without worksheet 10 – indeed, that might even be preferable, since the final section of the quarter often occurs a few days prior to the final.

In this worksheet, two data sets are presented, one (Data Set 1) that comes from growing yeast in an environment where resources are basically unlimited (the medium was changed every 3 hours) and another (Data Set 2) that comes from growing yeast in an environment where resources are limited (the medium was not changed). Data Set 1 is well-fit by an exponential, similar to the whooping crane data. Data Set 2 is well-fit by the logistic model (encountered in Worksheets 7 and 8). In this worksheet, the students estimate the one parameter in the exponential growth model (a) and two parameters from the logistic model (the growth rate, r, and carrying capacity, K) from Data Sets 1 and 2. They then compare their models to the data, using R.

The ideas of stability are introduced in estimating the parameter r in the logistic model. This is introduced intuitively, and is not explicitly discussed as linear stability analysis. The idea is for the students to work through this problem and then see the concept in lecture.

Things to do in lecture

This worksheet is aimed at motivating the discussion of stability of fixed points, so should (ideally) be introduced prior to section 5.6.

Materials

There are no materials for this discussion sheet.
Problem 1. Yeast population growth

We have looked at how populations grow in previous worksheets. Here, we will revisit population growth by examining how a population of yeast grows when resources are plentiful and when resources are scarce.

Nearly a century ago, Oscar Richards performed the following two experiments.

Experiment 1: Richards grew yeast in a container and changed the medium every three hours. In this way, he washed away the yeasts’ metabolism by-products and replenished their food. Over the course of the experiment, the yeast had, effectively, unlimited resources. Richards’ measurements from this experiment are given in the table below, in Table 13.

Experiment 2: Richards grew yeast in a container and did not change the medium. Here, the yeasts’ metabolism by-products built up and their food was depleted. In this experiment, the yeast had limited resources. Richards’ measurements from this experiment are given in the tables below, in Table 14.

<table>
<thead>
<tr>
<th>time (hours)</th>
<th>density of yeast</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>3.4</td>
</tr>
<tr>
<td>10</td>
<td>3.4</td>
</tr>
<tr>
<td>20</td>
<td>10.2</td>
</tr>
<tr>
<td>25</td>
<td>19.9</td>
</tr>
<tr>
<td>30</td>
<td>41.5</td>
</tr>
<tr>
<td>32.5</td>
<td>46.6</td>
</tr>
<tr>
<td>37.5</td>
<td>76.1</td>
</tr>
<tr>
<td>42.5</td>
<td>109.7</td>
</tr>
<tr>
<td>45</td>
<td>133.0</td>
</tr>
<tr>
<td>50</td>
<td>188.1</td>
</tr>
<tr>
<td>52.5</td>
<td>203.4</td>
</tr>
<tr>
<td>55</td>
<td>268.8</td>
</tr>
<tr>
<td>57.5</td>
<td>272.7</td>
</tr>
</tbody>
</table>

Table 13: Unlimited Resources (Exp.1)

<table>
<thead>
<tr>
<th>time (hours)</th>
<th>density of yeast</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>3.4</td>
</tr>
<tr>
<td>20</td>
<td>8.5</td>
</tr>
<tr>
<td>25</td>
<td>13.1</td>
</tr>
<tr>
<td>30</td>
<td>22.7</td>
</tr>
<tr>
<td>40</td>
<td>33.5</td>
</tr>
<tr>
<td>50</td>
<td>41.5</td>
</tr>
<tr>
<td>60</td>
<td>50.0</td>
</tr>
<tr>
<td>70</td>
<td>54.5</td>
</tr>
<tr>
<td>80</td>
<td>58.5</td>
</tr>
<tr>
<td>90</td>
<td>60.2</td>
</tr>
<tr>
<td>100</td>
<td>60.2</td>
</tr>
<tr>
<td>110</td>
<td>61.4</td>
</tr>
<tr>
<td>120</td>
<td>60.8</td>
</tr>
</tbody>
</table>

Table 14: Limited Resources (Exp.2)
1. Use R/R-studio to plot the data for each of the experiments. What differences do you notice between the two experiments? What similarities do you notice? Briefly explain.

2. When resources are unlimited, we might expect to see exponential population growth (as we saw with the Whooping Cranes earlier this quarter). In parts (a)–(d) below we will examine this hypothesis.

(a) Use R/R-studio to make a semi-log plot of the data in the left two columns of Table 1 (see sample code in part b, below).

(b) Fit the data with a straight line.

Recall R/R-studio sample code:
\[
t=c(0, 8.0, 10.0, 20.0, 25.0, 30.0, 32.5, 37.5, 42.5, 45.0, 50.0, 52.5, 55.0, 57.5)
\]
\[
N=c(2, 3.4, 3.4, 10.2, 19.9, 41.5, 46.6, 76.1, 109.7, 133.0, 188.1, 203.4, 268.8, 272.7)
\]
\[
F<-lm(log(N,base = 10) \sim t)
\]
\[
plot(t,log(N,base = 10))
\]
\[
lines(t,fitted(F))
\]

(c) Based on your fit, find a relationship between yeast density \(N\) and time \(t\).

(d) Rewrite your expression in part c) as a recursion of the form \(N_{i+1} = aN_i\) (where \(a\) is a constant you must determine), and \(N_i\) is the yeast density at hour \(i\).

3. When resources are limited, we don’t see exponential population growth. Instead, the population reaches a steady-state. Perhaps you have seen a model for resource-limited growth, described by the following recursion:

\[N_{i+1} = N_i + rN_i(1 - N_i/K)\]

where \(r\) and \(K\) are positive constants. Let’s see if this model describes Richards’ data with limited resources.

(a) Find all fixed points of the recursion.

(Recall: A fixed point, \(N^*\), of the recursion \(N_{i+1} = f(N_i)\) satisfies the equation \(N^* = f(N^*)\))

(b) Recall that, given a recursion for the sequence \(N_i\), the limit of the sequence is a fixed point. Which of the two fixed points you found in part (a) is the limit of the recursion?

(c) Looking at the data in Table 14, make a reasonable approximation for the limit. Use that result to determine one of the two parameters, \(r\) or \(K\).
(Hint: the parameter \( K \) is sometimes called the “carrying capacity”)

It is possible to model the initial behavior of both populations with the unlimited resources model, since initially the resources are not that limited. This idea will help us find an approximation for the other parameter in our model \((r \text{ or } K)\).

(d) What is the growth rate when the population is very small? To answer this, first linearize the model about \( N_i = 0 \). I.e. find a linear approximation of \( f(N_i) = N_i + rN_i(1 - N_i/K) \) near \( N_i = 0 \), and use this linear function as your linear recursion of the form \( N_{i+1} = bN_i \). (Note that when you do the linearization this will determine \( b \).)

(e) Now use the recursion from part 2(d) to determine the second parameter \((r \text{ or } K)\). Note that we have assumed that the recursion you found in part (d) is the same as the recursion you found in part 2(d).

4. Now you have two recursions that model the data in Tables 1 and 2. Use R/R-studio to make a table of data for your recursions and plot the recursions along with the original data (see sample code, below). Are your answers reasonable?

For example to make the limited growth recursion in R/R-studio, type:

```r
for (i in 1:120) {N[i+1]=N[i]+r*N[i]*(1-N[i]/K)}
```

(you will have to fill in \( r \) and \( K \) that you determine above)
Problems for week 9: stability of fixed points

1. **Solving a differential equation.**

Differential equations give a relationship between a function (say, \( x(t) \)) and its derivative (\( dx/dt \)). Consider, for example, the differential equation

\[
\frac{dx}{dt} = ax
\]  

(12)

where \( a \) is a constant (here, it can be either negative or positive). To “solve” this equation, we want to find the function \( x(t) \) that satisfies it. We can now solve it – and it turns out that the solution to this equation is very important (you’ll learn more about it in 17B).

To solve Eq. 12, let’s rewrite it in limit form:

\[
\lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} = ax(t)
\]  

(13)

Let’s now get rid of the limit (which gives us an approximation) and we have an equation that’s easier to work with

\[
\frac{x(t + \Delta t) - x(t)}{\Delta t} \approx ax(t)
\]  

(14)

**a)** Suppose that \( t = n\Delta t \) and that \( X_n = x(t) = x(n\Delta t) \). Write Eq. 14 as a recursion. To do so, you’ll have to specify an initial condition. Suppose that \( x(t = 0) = x_0 \), a constant.

**b)** Solve the recursion from part a.

**c)** Recall that \( t = n\Delta t \). In your answer from part b, use this equation to eliminate the index \( n \).

(i.e. suppose your answer to part b is \( X_n \approx 4n - 7 \). Then, since \( X_n = x(t) \) and since \( n = t/\Delta t \), your answer would be \( x(t) \approx 4t/\Delta t - 7 \).

**d)** Now, take the limit as \( \Delta t \to 0 \) to find an exact expression for \( x(t) \).

(HINT: your limit should be of the form \( 1^\infty \). Look at example 12 of the book, on page 251 if you are unsure how to solve this limit).

**e)** If you’ve done everything right, you now have an answer. Check your answer by showing that it satisfies the initial differential equation, Eq. 12.
2. Stability of fixed points in differential equations.

In this problem, we'll look at the fixed points of a differential equation and determine their stability, and compare and contrast to what we know about fixed points of difference equations.

The following differential equation is a model for resource-limited population growth,

\[ \frac{dN}{dt} = rN \left(1 - \frac{N}{K}\right) \]  

(15)

where \( r \) and \( K \) are positive constants.

We can convert it to a recursion using the same tricks as in problem 1:

\[
\lim_{\Delta t \to 0} \frac{N(t + \Delta t) - N(t)}{\Delta t} = rN(t) \left(1 - \frac{N(t)}{K}\right)
\]

\[
\frac{N_{i+1} - N_i}{\Delta t} \approx rN_i \left(1 - \frac{N_i}{K}\right)
\]

\[
N_{i+1} \approx r\Delta t N_i \left(1 - \frac{N_i}{K}\right) + N_i
\]

To define the recursion, we would also need the initial condition \( N(0) = N_0 \).

a) Find all fixed points of the recursion and show that, at each fixed point, \( \frac{dN}{dt} = 0 \) in Eq. 15.

b) In your answer to part a, one of your fixed points should be \( N^* = 0 \). The recursion is of the form \( N_{i+1} = f(N_i) \). Linearize the function \( f(N_i) \) about 0. Your answer should be similar to the recursion you found in problem 1a.

c) Since the recursion is the same as in problem 1, you should be able to infer from your solution to problem 1 what \( N(t) \) is near 0. Write down this expression.

d) What is required for the fixed point \( N^* = 0 \) to be stable?

e) From the recursion you found in part b), explain why you could never see oscillations in this model. (i.e. you could never get a recursion of the form \( N_{i+1} = cN_i \), where \( c < 0 \)).
Purpose and summary

This is a continuation of the ideas introduced in worksheet 9. It deals with stability of fixed points of difference equations. The idea is to emphasize how linearization is used to determine stability, and also to emphasize that linearization gives a good approximation of the solution in the neighborhood of the fixed point (and idea that becomes very important in 17C).

As in worksheet 9, two data sets are presented, one (Data Set 1) that comes from growing yeast in an environment where resources are basically unlimited (the medium was changed every 3 hours) and another (Data Set 2) that comes from growing yeast in an environment where resources are limited (the medium was not changed). Data Set 1 is well-fit by an exponential, similar to the whooping crane data. Data Set 2 is well-fit by the logistic model (encountered in Worksheets 7 and 8).

In this worksheet, the students approximate the solution near the unstable fixed point (0) and near the stable fixed point ($K$). To do so, they plot the first few points in Data Set 2 on a semi-log plot, to see it is linear. This gives an estimate of $r$ – and is similar to what they did on Worksheet 9. To approximate the solution near the stable fixed point, they plot $(N_i - K)$ on a semi-log plot, where $N_i$ is the population size and $K$ is the carrying capacity. This also gives an estimate of $r$, and is slightly different from what they found near the unstable fixed point. Finally, they plot the data along with the two linear approximations.

Things to do in lecture

This worksheet is aimed at emphasizing that linearization tells you more than just the stability of fixed points, but also gives an approximation for the solution near the fixed points. It should (ideally) be introduced after section 5.6.

Materials

There are no materials for this discussion sheet.
Week 10: Pre-discussion assignments
Problem 1. Yeast population growth, part II Recall the experiments of Oscar Richards, from the last worksheet.

Experiment 1: Richards grew yeast in a container and changed the medium every three hours. In this way, he washed away the yeasts’ metabolism by-products and replenished their food. Over the course of the experiment, the yeast had, effectively, unlimited resources. Richards’ measurements from this experiment are given in the table below, in the left-most two columns.

Experiment 2: Richards grew yeast in a container and did not change the medium. Here, the yeasts’ metabolism by-products built up and their food was depleted. In this experiment, the yeast had limited resources. Richards’ measurements from this experiment are given in the table below, in the right-most two columns.

Table 15: Data from O. W. Richards, 1928

<table>
<thead>
<tr>
<th>time (hours)</th>
<th>density of yeast</th>
<th>time (hours)</th>
<th>density of yeast</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>unlimited resources</td>
<td></td>
<td>limited resources</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>3.4</td>
<td>10</td>
<td>3.4</td>
</tr>
<tr>
<td>10</td>
<td>3.4</td>
<td>20</td>
<td>8.5</td>
</tr>
<tr>
<td>20</td>
<td>10.2</td>
<td>25</td>
<td>13.1</td>
</tr>
<tr>
<td>25</td>
<td>19.9</td>
<td>30</td>
<td>22.7</td>
</tr>
<tr>
<td>30</td>
<td>41.5</td>
<td>40</td>
<td>33.5</td>
</tr>
<tr>
<td>32.5</td>
<td>46.6</td>
<td>50</td>
<td>41.5</td>
</tr>
<tr>
<td>37.5</td>
<td>76.1</td>
<td>60</td>
<td>50.0</td>
</tr>
<tr>
<td>42.5</td>
<td>109.7</td>
<td>70</td>
<td>54.5</td>
</tr>
<tr>
<td>45</td>
<td>133.0</td>
<td>80</td>
<td>58.5</td>
</tr>
<tr>
<td>50</td>
<td>188.1</td>
<td>90</td>
<td>60.2</td>
</tr>
<tr>
<td>52.5</td>
<td>203.4</td>
<td>100</td>
<td>60.2</td>
</tr>
<tr>
<td>55</td>
<td>268.8</td>
<td>110</td>
<td>61.4</td>
</tr>
<tr>
<td>57.5</td>
<td>272.7</td>
<td>120</td>
<td>60.8</td>
</tr>
</tbody>
</table>
Recall that, when resources are limited, the population was reasonably described by the following recursion

\[ N_{i+1} = N_i + rN_i(1 - N_i/K) \]

where \( r \) and \( K \) are positive constants.

Last time, you saw that \( N^* = 0 \) and \( N^* = K \) were the two fixed points of the recursion.

(recall: a fixed point, \( N^* \), of the recursion \( N_{i+1} = f(N_i) \) satisfies the equation \( N^* = f(N^*) \))

a) Determine the stability of the fixed point at \( N^* = 0 \). Does your answer make sense?

b) Recall that stability is based on the idea of linearization. Linearize the non-linear recursion \( (N_{i+1} = N_i + rN_i(1 - N_i/K)) \) about the fixed point \( N^* = 0 \), and solve the recursion. (This is similar to what you did on the last worksheet)

c) The solution to a recursion is an exponential, so that the data should be linear near the fixed point \( N^* = 0 \), if plotted appropriately. Use R/R-studio to create the appropriate plot, and determine the best-fit line.

HINTS: You don’t want to fit all of the data, just the first four or five points. See the sample R code on the last page of the worksheet for ideas about how to do this.

d) Determine the stability of the fixed point at \( N^* = K \). Does your answer make sense?

e) Recall that stability is based on the idea of linearization. Linearize the non-linear recursion \( (N_{i+1} = N_i + rN_i(1 - N_i/K)) \) about the fixed point \( N^* = 0 \), and solve the recursion.

f) The solution to a recursion is an exponential, so that the data should be linear near the fixed point \( N^* = K \), if plotted appropriately. Use R/R-studio to create the appropriate plot, and determine the best-fit line.

HINTS: You don’t want to fit all of the data, just the last four or five points. You’ll need to decide on a numerical value for \( K \), the carrying capacity. You will likely get an error when you try to take the log of the data, because some of the data will be above the carrying capacity. You want to ignore these points in your linear fit. See the sample R code on the last page of the worksheet for ideas about how to do this.

g) In parts c and f, you should obtain estimates of the parameter \( r \). It is likely that they do not precisely agree. Which one do you trust more? Why?
h) Using R/R-studio, plot the data along with the two approximations.

**Sample R code** that will plot all of the data from experiment 2 on a semi-log plot and then fit it with a straight line. This is meant to be a starting point for your code; you’ll need to write your own code to complete the worksheet.

t=c(0, 10.0, 20.0, 25.0, 30.0, 40.0, 50.0, 60.0, 70.0, 80.0, 90.0, 100.0, 110.0, 120.0)
N=c(2, 3.4, 8.5, 13.1, 22.7, 33.5, 41.5, 50.0, 54.5, 58.5, 60.2, 60.2, 61.4, 60.8)
F<-lm(log(N,base = 10)∼t)
plot(t,log(N,base = 10))
lines(t,fitted(F))
Problems for week 10: stability of fixed points

See problems for week 9.