Section notes for POLSCI 598:
Mathematics for Political Science

Samuel Baltz
sbaltz@umich.edu
Fall 2020

Written to complement lectures by Iain Osgood

Contents

1 Section 1: 2020 September 22

Topics:

- Introductory details[1]
- Derivatives of exponentials[2]
- Derivatives of logarithms[3]
- An application[4]

After this class, I expect every student to be able to:

- Explain the importance of e
- Follow the steps to compute a 2×2 Hessian

BACK TO TABLE OF CONTENTS

① **Introductory details**:

↪Drop-in Office Hours on Zoom (sigh) for 2 hours a week, we can use the same call that the class is in to keep things simple. You can tune in from 2306 MH if you want to.
↪I always say yes to 1 on 1 appointments, but please try to come to office hours first. If you can't make office hours, though, don't let that stop you from asking for an appointment.
↪I try to have grades by the following class, and grades will always be given within a week (for both problem sets and tests).
↪The only real way to reach me is emailing sbaltz@umich.edu. I nearly always respond to emails within 24 hours.
↪Typed section notes will be uploaded every Tuesday night summarising pretty much all of the information from that day's section.

② **Derivatives of exponentials**:

Let's begin our discussion with the following rule, which we briefly saw in math camp:

Recall: In order to take the derivative of some variable $x \in \mathbb{R}$ when it is the exponent of some constant $r \in \mathbb{R}$, we can use the rule:

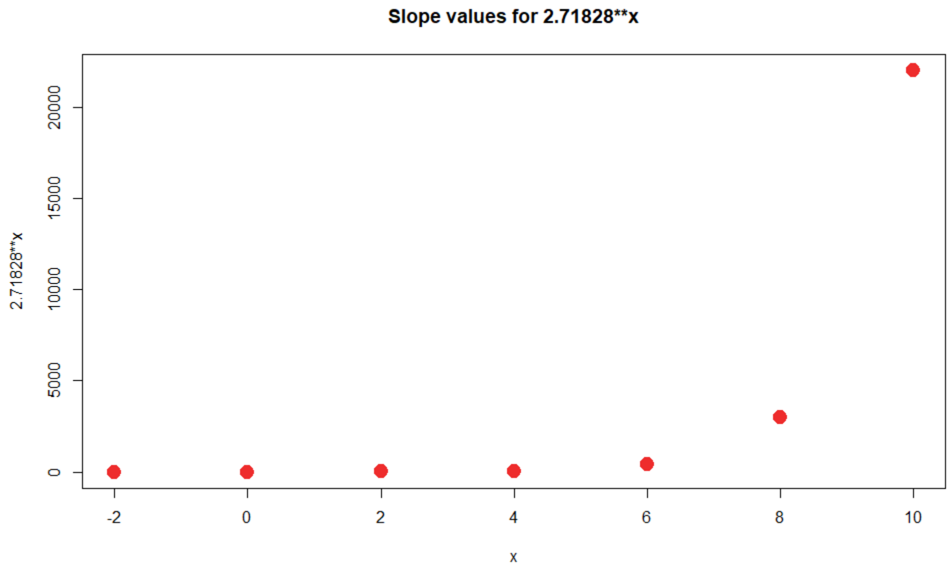
$$\frac{d}{dx}(r^x) = r^x \ln(r)$$

We're just going to take this rule as a given so that we can figure out what's so special about Euler's constant e. This interesting special property of e is that it satisfies the much simpler rule:

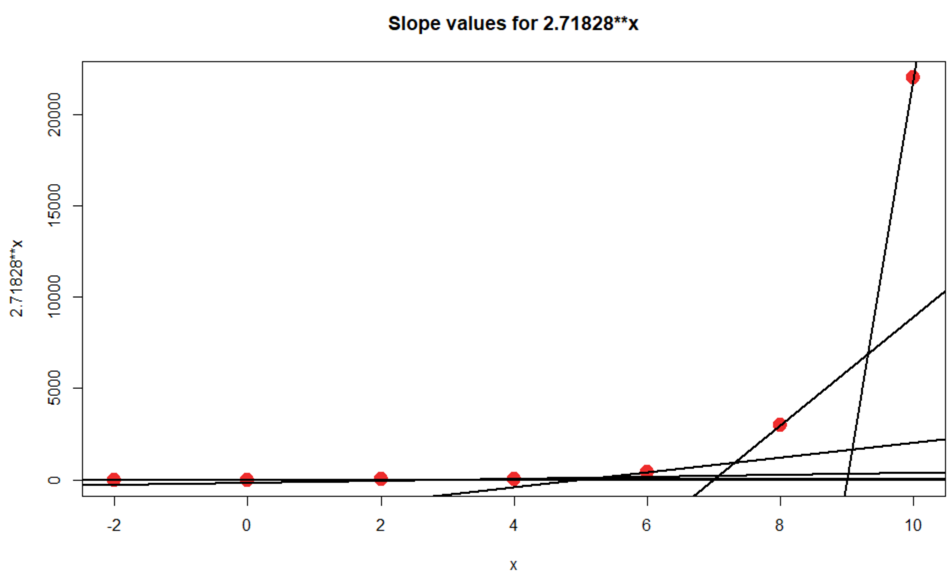
$$\frac{d}{dx}(e^x) = e^x$$

To see exactly what this means, I wrote some code in R (available on Canvas in the Section Materials folder)[1] that first asks you to choose a constant r, and let's cut to the chase by choosing $r = e$ (In section we played around with many different values for r to show that e is special; you can replicate this if you like by downloading the code and changing the value in the first line to something other than e.). The code does the following: first it plots r^x against r in any range you choose, like this:

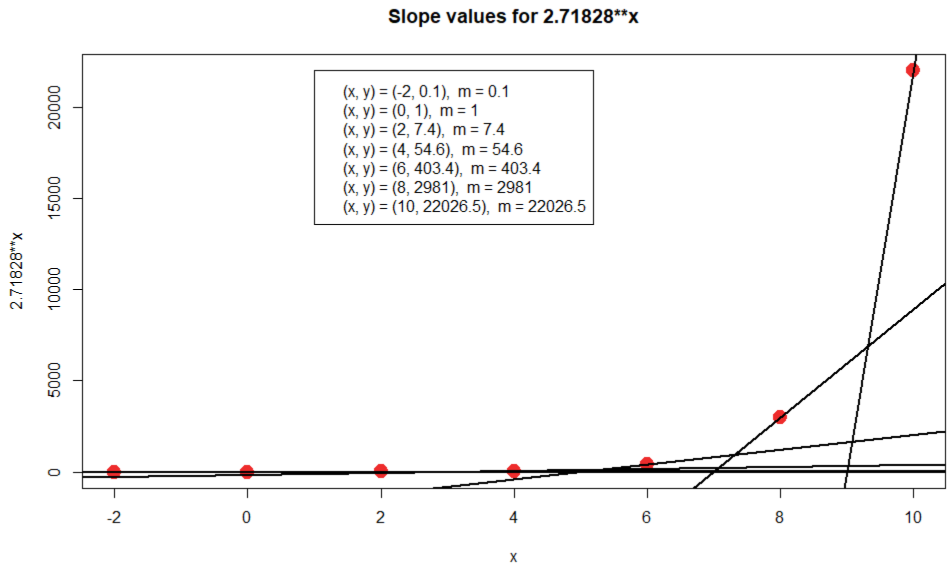
[1] The full code for any computer program that I show during sections will always be posted to Canvas so that you can do whatever you want with it.



Next it calculates the derivative of the curve at each of those points, using our rule for taking a derivative with respect to an exponent. By the derivative of the curve at each point, we mean the line that lies tangent to the curve at that point. Drawing these lines, we get:



Now let's write down the (x, y) values of the points beside the value of the slope m of the tangent at each point:



The quite amazing thing about this plot is that the m values exactly match the y-values of the corresponding point of the function!

Remark: We know that at each point, $m = \frac{d}{dx}e^x$. Recalling that the y-values are given by $y = e^x$, the fact that the slope values exactly match the y-values therefore means that

$$m = e^x$$

So,

$$\frac{d}{dx}e^x = e^x$$

If you play around with the R code, you will see that e is quite special in this regard. Changing the constant r so that it is more distant from e will also change the slope values so that they are more distant from the corresponding y-values.

Note: This result checks out with our rule for taking a derivative with respect to an exponent. Recall that $\ln(e) = 1$. Using the rule,

$$\frac{d}{dx}e^x = e^x \ln(e)$$

$$\frac{d}{dx}e^x = e^x \cdot 1$$

$$\frac{d}{dx}e^x = e^x$$

③ **Derivatives of logarithms**:

Using this property of e, we can figure out another tricky fact: $\frac{d}{dx}\ln(x)$. Let's make things easier on ourselves by assuming that $\ln(x)$ is differentiable (this is far from being obviously true, but it turns out to be fine). We're also going to need the following rule:

Recall: The **chain rule** says that the derivative of a composition of functions $f \circ g$ is the derivative of the outside function times the derivative of the inside function, so

$$(f \circ g)' = (f' \circ g) \cdot g'$$

Now we can begin with the fact

$$\frac{d}{dx}x = 1$$

Since $e^{\ln(x)} = x$ by definition, we can substitute in

$$\frac{d}{dx}e^{\ln(x)} = 1$$

Now let's try to figure out what is the derivative of $e^{\ln(x)}$. Our assumption that $\ln(x)$ is differentiable allows us to apply the chain rule. So, using our special property of e,

$$\frac{d}{dx}e^{\ln(x)} = e^{\ln(x)}\frac{d}{dx}\big(\ln(x)\big)$$

With $e^{\ln(x)} = x$ by definition,

$$\frac{d}{dx}e^{\ln(x)} = x\frac{d}{dx}\big(\ln(x)\big)$$

Now, using our identity above, substitute

$$1 = x\frac{d}{dx}\big(\ln(x)\big)$$

Dividing both sides by x,

$$\frac{d}{dx}\big(\ln(x)\big) = \frac{1}{x}$$

□

④ **An application**:

Let's do a question that is very similar to the first question on Problem Set 2. First let's remind ourselves of another incredibly important rule:

Recall: The **product rule** says that the derivative of the product of any two differentiable functions f and g is given by:

$$(f \cdot g)' = f' \cdot g + g' \cdot f$$

Question: Find the Hessian of $f(x, y) = e^{xy+1}$

Answer: Recall that the Hessian matrix of a function $f(x, y)$ is given by

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

so to compute the Hessian we should compute each element individually. Let's go through each one.

$$\frac{\partial^2}{\partial x^2} f(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\left(e^{xy+1}\right)\right)$$

$$\frac{\partial^2}{\partial x^2} f(x, y) = \frac{\partial}{\partial x}\left(e^{xy+1}\frac{\partial}{\partial x}\left(xy+1\right)\right)$$

$$\frac{\partial^2}{\partial x^2} f(x, y) = \frac{\partial}{\partial x}\left(ye^{xy+1}\right)$$

By the product rule,

$$\frac{\partial^2}{\partial x^2} f(x, y) = e^{xy+1}\frac{\partial}{\partial x}(y) + y\frac{\partial}{\partial x}\left(e^{xy+1}\right)$$

$$\frac{\partial^2}{\partial x^2} f(x, y) = y^2 e^{xy+1}$$

Now let's do the cross partial derivatives:

$$\frac{\partial^2}{\partial y \partial x} f(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}\left(e^{xy+1}\right)\right)$$

We just calculated the x derivative so we can plug in that result directly:

$$\frac{\partial^2}{\partial y \partial x} f(x, y) = \frac{\partial}{\partial y}\left(ye^{xy+1}\right)$$

then by the product rule,

$$\frac{\partial^2}{\partial y \partial x} f(x, y) = e^{xy+1} \frac{\partial}{\partial y}(y) + y \frac{\partial}{\partial y}\left(e^{xy+1}\right)$$

$$\frac{\partial^2}{\partial y \partial x} f(x, y) = e^{xy+1} + yxe^{xy+1}$$

$$\frac{\partial^2}{\partial y \partial x} f(x, y) = e^{xy+1}(xy + 1)$$

And recall that

$$\frac{\partial^2}{\partial y \partial x} f(x, y) = \frac{\partial^2}{\partial x \partial y} f(x, y)$$

Finally,

$$\frac{\partial^2}{\partial y^2} f(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial y}\left(e^{xy+1}\right)\right)$$

$$\frac{\partial^2}{\partial y^2} f(x, y) = \frac{\partial}{\partial y}\left(e^{xy+1} \frac{\partial}{\partial y}(xy + 1)\right)$$

$$\frac{\partial^2}{\partial y^2} f(x, y) = \frac{\partial}{\partial y}\left(xe^{xy+1}\right)$$

By the product rule,

$$\frac{\partial^2}{\partial y^2} f(x, y) = e^{xy+1} \frac{\partial}{\partial y}(x) + y \frac{\partial}{\partial y}\left(e^{xy+1}\right)$$

$$\frac{\partial^2}{\partial y^2} f(x, y) = x^2 e^{xy+1}$$

Note: We could have done this without repeating any work just by observing that the variables x and y can be swapped without loss of generality in the original function. That is, because x and y appear in identical ways in the function (simply renaming x to y and y to x would yield an identical function), we know that the second partial with respect to y is going to mirror the second partial with respect to x.
Finally, we can plug these values into the Hessian:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} y^2 e^{xy+1} & e^{xy+1}(xy+1) \\ e^{xy+1}(xy+1) & x^2 e^{xy+1} \end{bmatrix}$$

2 Section 2: 2020 September 29

Topics:

- Taylor series[1]
- First and second directional derivatives[2]

After this class, I expect every student to be able to:

- Take the n^{th} order Taylor series approximations of simple functions for small n
- Take first and second directional derivatives of a simple function

BACK TO TABLE OF CONTENTS

① Single-variable Taylor series approximations:

Recall: In lecture we saw that the Taylor Series at point a of a function $f(x)$ that can be differentiated any number of times at point a is given by

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots$$

Note: This is much easier to remember if you notice that the first term is not an exception to the rule, because the simplified (and most common) way of writing the Taylor Series is exactly the same as writing:

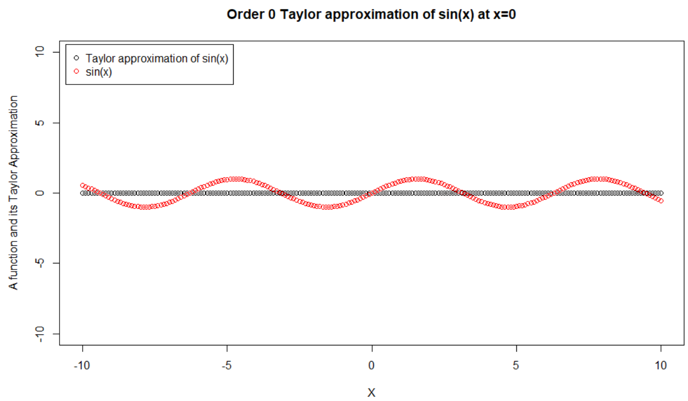
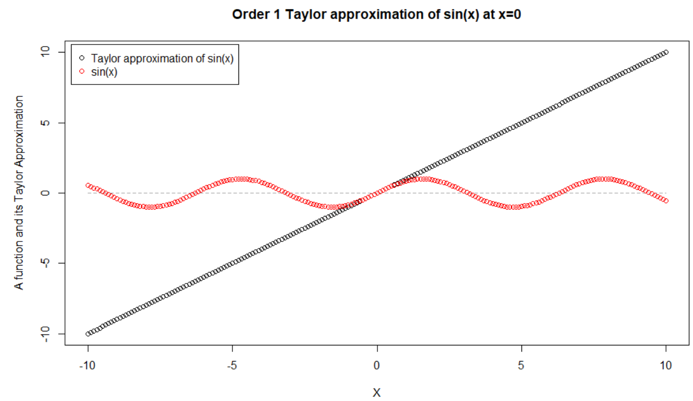
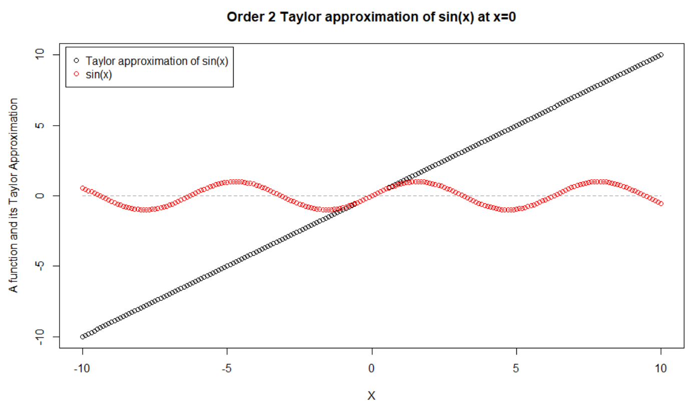
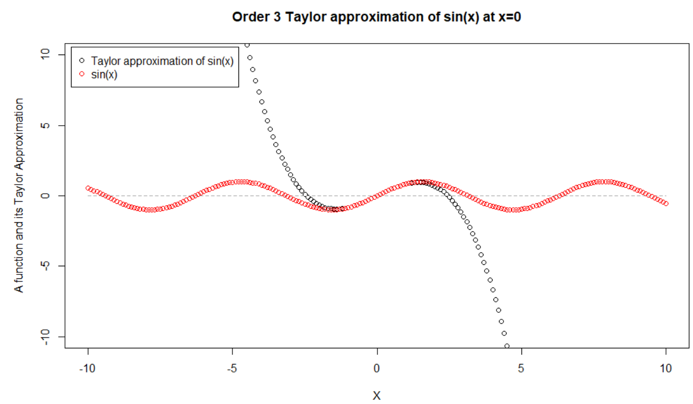
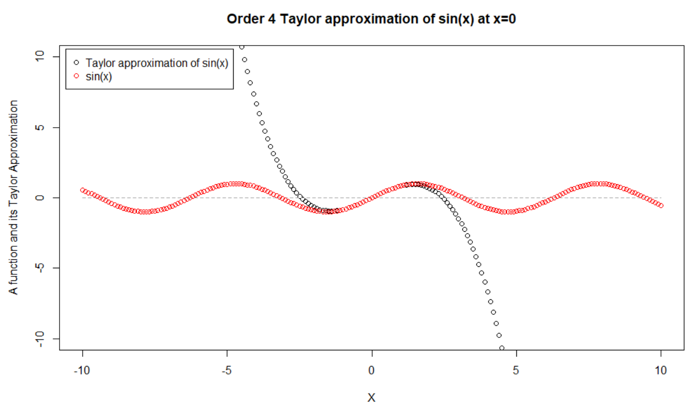
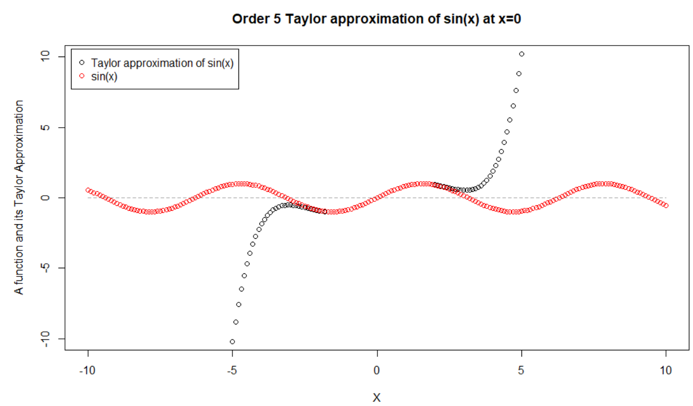
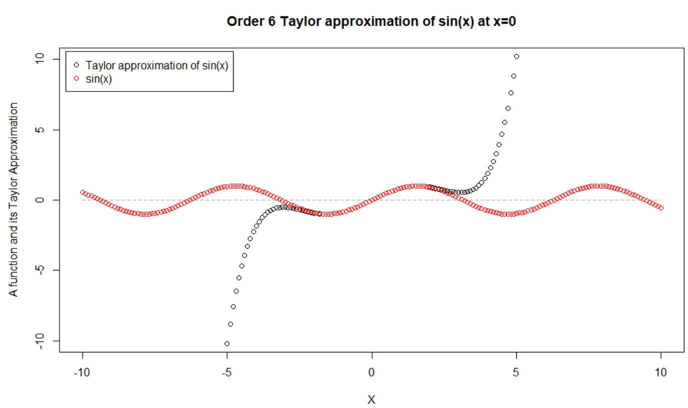
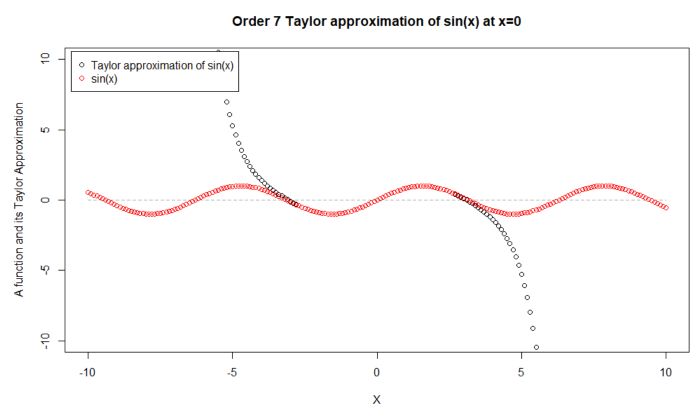
$$f(x) = \frac{f(a)}{0!}(x-a)^0 + \frac{f'(a)}{1!}(x-a)^1 + \frac{f''(a)}{2!}(x-a)^2 + \cdots$$

In fact, to remember the Taylor Series, I always think to myself "Take the derivative zero times, divide by zero factorial, and multiply by $(x-a)$ to the power of zero. In the next term take the derivative once, divide by one factorial, and multiply by $(x-a)$ to the power of one. Then, ..." With this thought process, you only need to remember the three pieces that go into all of the terms.

Recall: When Iain says or writes, for example, "a Taylor approximation of order 2", recall that the "order" of a polynomial (or the degree of a polynomial) usually means the largest exponent of that polynomial. So a Taylor approximation of order 2 actually means up to the 3rd term – that is, the 2^{nd} order Taylor approximation of $f(x)$ which is twice differentiable at point a is

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2$$

Exercise: Calculate the n^{th} order Taylor approximations of $\sin(x)$ for increasing values of n, starting at 0, and centering at first on point $a = 0$. We did a few of these, and the first 8 are shown in the following figure.


Order 0 Taylor approximation of sin(x) at x=0
Taylor approximation of sin(x)
sin(x)
A function and its Taylor Approximation
X

Order 1 Taylor approximation of sin(x) at x=0
Taylor approximation of sin(x)
sin(x)
A function and its Taylor Approximation
X

Order 2 Taylor approximation of sin(x) at x=0
Taylor approximation of sin(x)
sin(x)
A function and its Taylor Approximation
X

Order 3 Taylor approximation of sin(x) at x=0
Taylor approximation of sin(x)
sin(x)
A function and its Taylor Approximation
X

Order 4 Taylor approximation of sin(x) at x=0
Taylor approximation of sin(x)
sin(x)
A function and its Taylor Approximation
X

Order 5 Taylor approximation of sin(x) at x=0
Taylor approximation of sin(x)
sin(x)
A function and its Taylor Approximation
X

Order 6 Taylor approximation of sin(x) at x=0
Taylor approximation of sin(x)
sin(x)
A function and its Taylor Approximation
X

Order 7 Taylor approximation of sin(x) at x=0
Taylor approximation of sin(x)
sin(x)
A function and its Taylor Approximation
X

Exercise: This figure has a very surprising and mysterious property: we discovered that with $a = 0$, all of the even approximations of $\sin(x)$ (beyond order 0) are no better than the preceding approximation, so for example the order 2 Taylor approximation to $\sin(x)$ is no better than the order 1 approximation to $\sin(x)$. Also, we would probably expect the order 2 approximation to look quadratic, and it really does not. To try to understand this, we calculated the order 2 and order 4 Taylor approximation to $\sin(x)$. The order two term in the Taylor approximation to $\sin(x)$, call this term t_2, is given by:

$$t_2 = \frac{\sin''(a)}{2!}(x-a)^2$$

So with $a = 0$,

$$t_2 = \frac{-\sin(0)}{2!}(x-0)^2$$

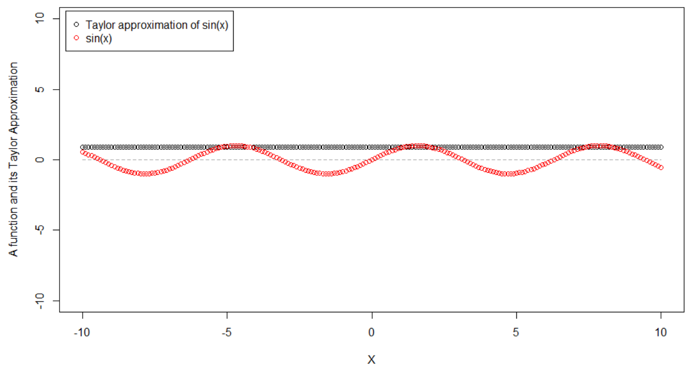
$$t_2 = 0$$

The order 4 approximation is also 0 for similar reasons. So it makes sense that it does not change the shape of the approximation at all or make it any more accurate; if we take the derivative of $\sin(x)$ an even number of times, we will always obtain either $\sin(x)$ or $-\sin(x)$, both of which evaluate to 0 when x is a multiple of 2π. But odd order approximations will evaluate $\cos(0)$ instead of $\sin(0)$, so they will obtain 1 instead of 0 and be nonzero functions of the expected order.

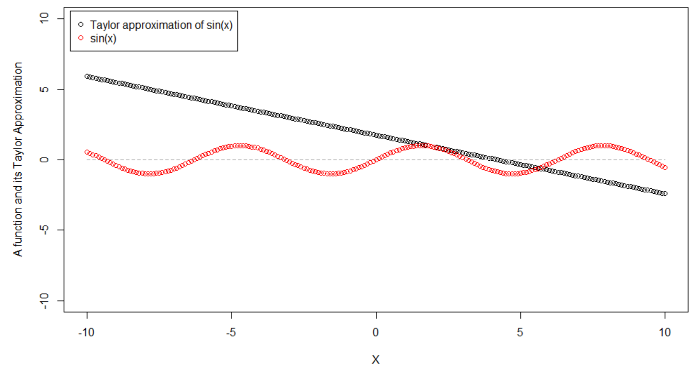
Note: We have discussed a few times in lecture that adding more terms onto the Taylor approximation does not *necessarily* make it more accurate, except in the limit of adding infinite terms onto it. This discovery about $\sin(x)$ gives a very natural and common example of that.

Part of this argument was that $a = 0$ (or, really, that $a \equiv 0 \mod 2\pi$). So what if we choose some other a, say $a = 2$? Those plots follow, and they also illustrate the broader points that a is the point at which the Taylor approximation always fits best, and also that the choice of a strongly affects the shape of the Taylor approximations.

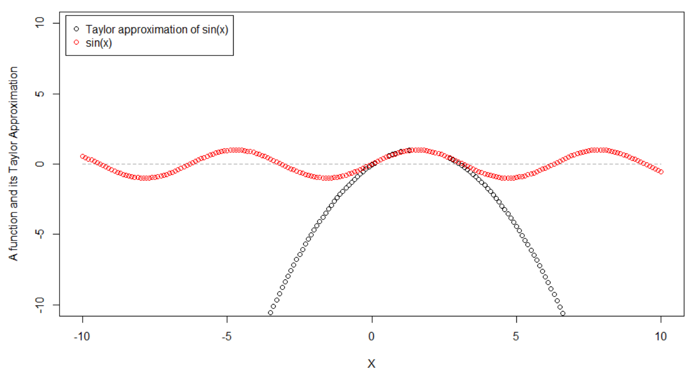
Order 0 Taylor approximation of sin(x) at x=2



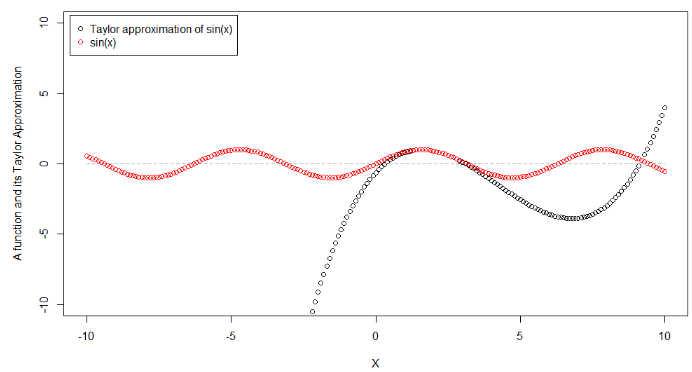
Order 1 Taylor approximation of sin(x) at x=2



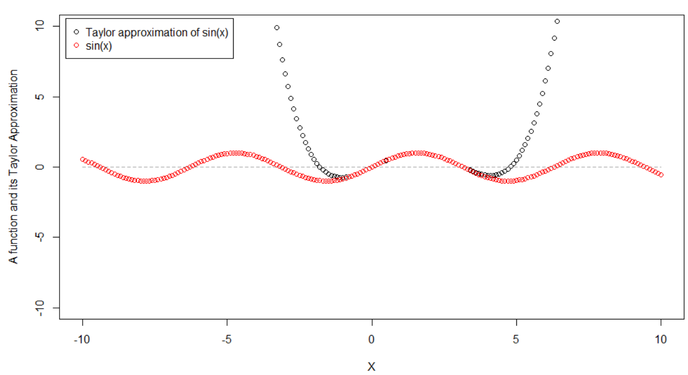
Order 2 Taylor approximation of sin(x) at x=2



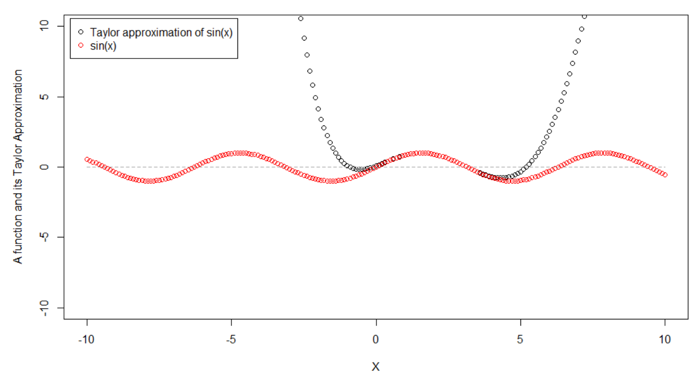
Order 3 Taylor approximation of sin(x) at x=2



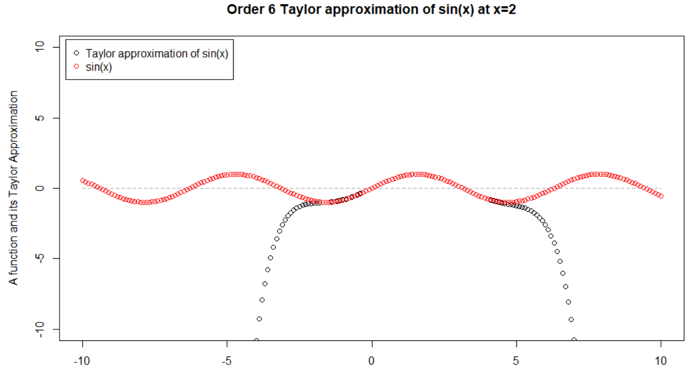
Order 4 Taylor approximation of sin(x) at x=2



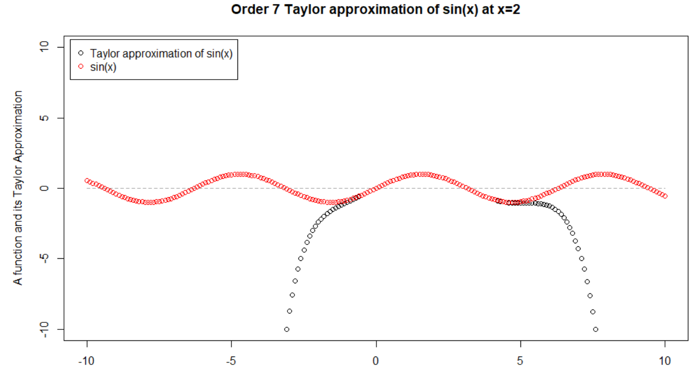
Order 5 Taylor approximation of sin(x) at x=2



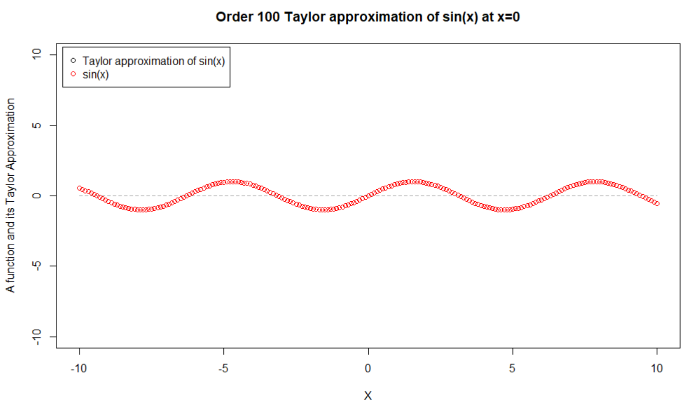
Order 6 Taylor approximation of sin(x) at x=2

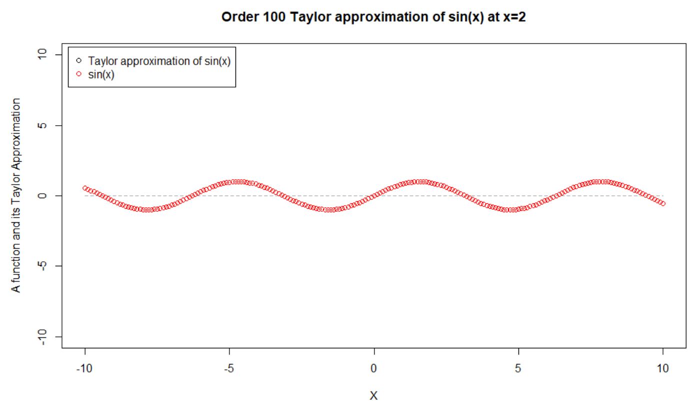


Order 7 Taylor approximation of sin(x) at x=2



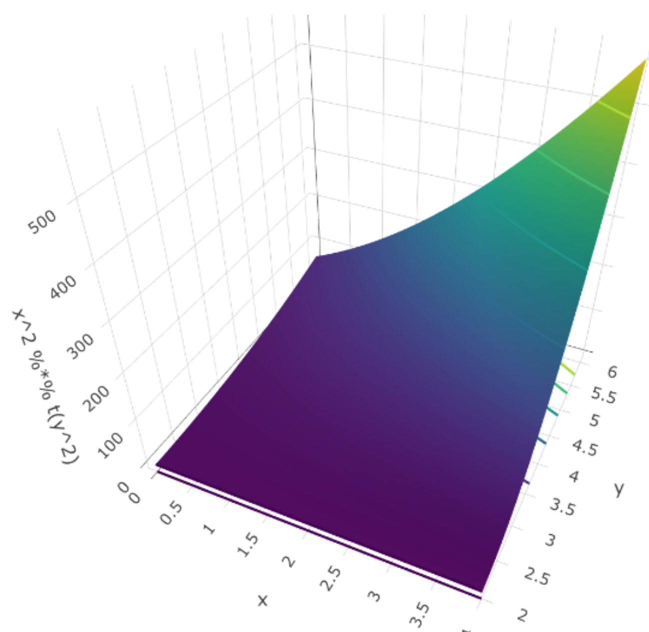
Next we quickly convinced ourselves that, no matter what a we choose, if we pick a high enough degree the Taylor fit will be really good:





② **Directional derivatives and Hessians**:

Exercise: Let's dive deeper on a function that we saw in an example in lecture: let's consider $f(x, y) = x^2y^2$. A 3D plot of this function on the range $x \in \{0; 4\}, y \in \{2; 6\}$ looks like this:



First, everyone individually computed the Hessian matrix of $f(x, y) = x^2y^2$, which is

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

$$H(f) = \begin{bmatrix} 2y^2 & 4xy \\ \\ 4xy & 2x^2 \end{bmatrix}$$

Let's first focus on the point $(a, b) = (2, 4)$. What is the Hessian at that point?

$$H(2,4) = \begin{bmatrix} 2 \cdot 4^2 & 4 \cdot 2 \cdot 4 \\ \\ 4 \cdot 2 \cdot 4 & 2 \cdot 2^2 \end{bmatrix}$$

$$H(2,4) = \begin{bmatrix} 32 & 32 \\ 32 & 8 \end{bmatrix}$$

Exercise: Now let's calculate the directional derivative at $(2, 4)$ looking to the point $(3, 3)$, as in the problem set. **Answer:** With reference to the lecture notes, the first directional derivative is given by

$$D_{(3,3)} f(2,4) = Df(2,4) \cdot \frac{u}{|u|}$$

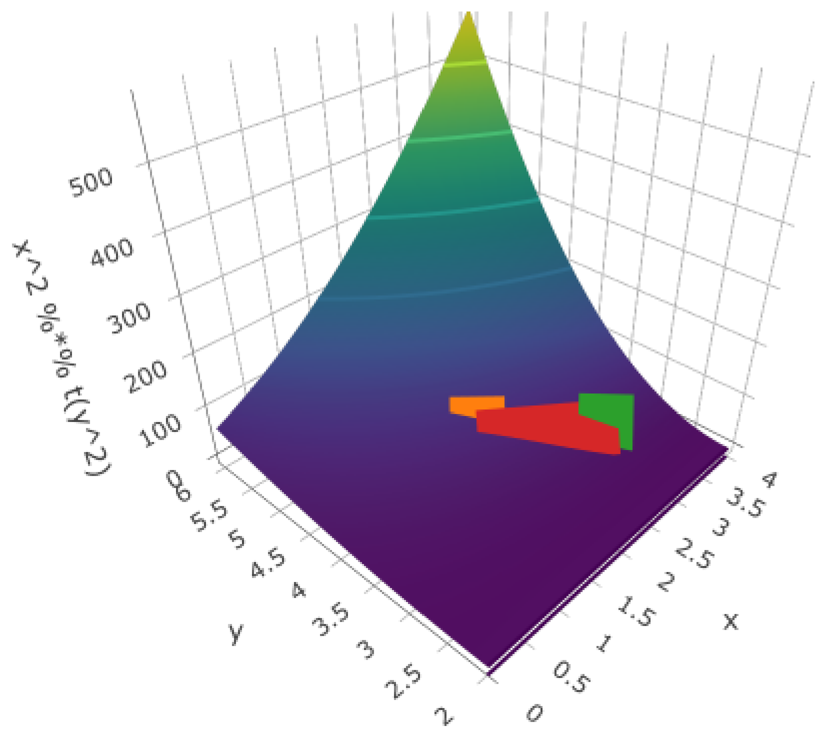
$$D_{(3,3)} f(2,4) = \begin{bmatrix} \frac{\partial f(2,4)}{\partial x} & \frac{\partial f(2,4)}{\partial y} \end{bmatrix} \cdot \frac{\begin{bmatrix} 3-2 & 3-4 \end{bmatrix}}{\sqrt{(3-2)^2 + (3-4)^2}}$$

$$D_{(3,3)} f(2,4) = \begin{bmatrix} 2xy^2\big|_{(2,4)} & 2x^2y\big|_{(2,4)} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$D_{(3,3)} f(2,4) = \left(2 \cdot 2 \cdot 4^2 - 2 \cdot 4 \cdot 4\right) \frac{1}{\sqrt{2}}$$

$$D_{(3,3)} f(2,4) \approx 22.63$$

This is the red line in the following plot, but if it were normalized to be of length 1, connecting (2,4) represented by the orange block to (3,3) represented by the green block:



Exercise: Because we're overachievers, let's also take the second directional derivative.
Answer: With reference to the lecture notes, the second directional derivative is given by

$$D^2_{(3,3)}f(2,4) = \frac{u}{|u|}H(2,4)\frac{u^T}{|u|}$$

$$D^2_{(3,3)}f(2,4) = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 32 & 32 \\ 32 & 8 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$D^2_{(3,3)}f(2,4) = -12$$

Note: Once upon a time, as an energetic and eager undergraduate, I struggled mightily trying to calculate directional derivatives in Wolfram|Alpha – no matter how many times I checked and re-checked my procedure, I couldn't get the numbers to match the computer's. The reason is this. To get a directional derivative in Wolfram|Alpha, for example the one we're currently studying, you can phrase the query as

```
derivative x^2y^2 at (2,4) in the direction (1,-1)
```

where what comes after “in the direction” is the vector direction, not the point that you are moving in the direction of. If you instead input

```
derivative x^2y^2 at (2,4) in the direction (3,3)
```

you will get the answer to a completely different question: Wolfram|Alpha thinks you’re talking about the vector $\begin{bmatrix} 3 & 3 \end{bmatrix}$, not the point $(3, 3)$. I once spent many miserable hours struggling with this, so please don’t waste your precious time on Earth the same way I did.

3 Section 3: 2020 October 6

Topics:

- Gauss-Jordan elimination using matrices[1]
- LU factorisation[2]

After this class, I expect every student to be able to:

- Perform Gauss-Jordan elimination using matrix multiplication as the elementary row operations
- Describe in words what the LU factorisation looks like

BACK TO TABLE OF CONTENTS

① **Gauss-Jordan elimination using matrices**:

Recall: A matrix is in **row-echelon form** if all of the following are true:
- ↪ The leftmost nonzero entry in each row (called the **leading entry**) is a 1,
- ↪ Every leading entry is to the right of the leading entry above it,
- ↪ Rows that only contain zeroes are at the bottom of the matrix.

Recall: A matrix is in **reduced row-echelon form** if
- ↪ It is in row-echelon form
- ↪ Each leading 1 is the only nonzero entry in its column

Recall: The **elementary row operations** are as follows:
- ↪ Switch two rows
- ↪ Multiply a row by a nonzero scalar
- ↪ Add a multiple of one row to another

Theorem 1. *Let $A \in \mathbb{R}^{m\times m}$ be a square matrix, $m \in \mathbb{N}$. A can be transformed into an $m\times m$ upper triangular matrix by left-multiplying A by a sequence of lower-triangular matrices L_k, as*

$$L_{m-1}\cdots L_2 L_1 A = U$$

Denoting the product $L_{m-1}\cdots L_2 L_1$ as L^{-1}, the matrix LU is a factorization of A.

Question: What is the connection between row-echelon form and upper triangularity?

Note: Theorem 1 describes an amazingly deep connection between the idea of Gauss-Jordan elimination and matrix multiplication. It is one example of how we can reduce a matrix to row-echelon form using matrix multiplication. This is built on the following incredible fact: each elementary row operation is equivalent to left-multiplication by a corresponding matrix. Those equivalencies are as follows:

Note: LU is a mnemonic: L stands for **L**ower triangular and U stands for **U**pper triangular. Nice!

↪In order to switch rows i and j, we can left-multiply by a slight modification of the identity matrix, where we have switched rows i and j. To switch rows 2 and 3 of a 3×3 matrix, we would use:

$$\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}=\begin{bmatrix}1&2&3\\7&8&9\\4&5&6\end{bmatrix}$$

$\hookrightarrow$In order to multiply row i by a nonzero scalar, we can left-multiply by another modification of the identity matrix, where we have multiplied row i by the scalar r. To multiply row 1 of a 3×3 matrix by 2, we would use:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Exercise: See if you can figure out the matrix that corresponds to the last elementary row operation: how would you multiply a scalar r times row i of a 3×3 matrix and add it to row j of that matrix? **Answer:**

$\hookrightarrow$In order to multiply row i by a nonzero scalar and add it to row j, we can left-multiply by another modification of the identity matrix, where we have replaced index j, i with the scalar r. So if we wanted to multiply 2 times row 2 to row 3 of our 3×3 matrix, we would use:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 \cdot 4 + 7 & 2 \cdot 5 + 8 & 2 \cdot 6 + 9 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 15 & 18 & 21 \end{bmatrix}$$

Some practice questions: Use the example matrix

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$\hookrightarrow$ Instead of switching two rows of $A_{2 \times 2}$, how would you switch two columns? Does this represent a legitimate row reduction?

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$$

Answer: Instead of left-multiplying by the row switching matrix, simply right-multiply:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$$

$\hookrightarrow$ Try to add 3 times row 2 to row 1, and also switch the rows of the matrix, using two matrix multiplications.

Answer:

$$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 2\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}7 & 10\\3 & 4\end{bmatrix}$$

$$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 2\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} = \begin{bmatrix}3 & 4\\7 & 10\end{bmatrix}$$

$\hookrightarrow$ **The Truth Hurts Challenge**: Now try to do the same thing using just one matrix multiplication. Extra challenge: Try to add 3 times row 2 to row 1, and also switch the *columns* of the matrix, with multiplication by just one 2×2 matrix.

Only the first can be done by just condensing the two left-multiplications into one, obtaining:

$$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 2\\0 & 1\end{bmatrix} = \begin{bmatrix}0 & 1\\1 & 2\end{bmatrix}$$

(2) **LU factorisation**:

Example: [2] We will use very similar ideas (although not exactly these matrices) to bring a matrix into its LU factorization. We will begin with a 4×4 matrix, and the idea will be to left-multiply it by 3 matrices, where each matrix converts another column into row-echelon form. Consider the matrix A defined as

$$A = \begin{bmatrix}2 & 1 & 1 & 0\\4 & 3 & 3 & 1\\8 & 7 & 9 & 5\\6 & 7 & 9 & 8\end{bmatrix}$$

First we want to turn every entry in the first column into a 0, except the entry in row 1. How are we going to do that? Surely we want to subtract twice row 1 from row 2, four times row 1 from row 3, and 3 times row 1 from row 4. With reference to our matrix that adds a multiple of one row to another row, it looks like this should do the job:

$$L_1 = \begin{bmatrix}1 & 0 & 0 & 0\\-2 & 1 & 0 & 0\\-4 & 0 & 1 & 0\\-3 & 0 & 0 & 1\end{bmatrix}$$

Checking that this works, then:

$$\begin{bmatrix}1 & 0 & 0 & 0\\-2 & 1 & 0 & 0\\-4 & 0 & 1 & 0\\-3 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}2 & 1 & 1 & 0\\4 & 3 & 3 & 1\\8 & 7 & 9 & 5\\6 & 7 & 9 & 8\end{bmatrix} = \begin{bmatrix}2 & 1 & 1 & 0\\0 & 1 & 1 & 1\\0 & 3 & 5 & 5\\0 & 4 & 6 & 8\end{bmatrix}$$

[2] This example is dramatically expanded from an example in (Trefethen and Bau III, 1997: p. 148), although for the parts that they covered, I tried to change as little as possible; if Trefethen and Bau III have already explained something, then it's already been explained just about as well as it can possibly be.

Our next move is to get rid of the 3 and 4 in column 2. It looks straightforward enough to do this: we can use the one in position (2,2) to subtract 3 times row 2 from row 3 and 4 times row 2 from row 4. The matrix that will do this for us is as follows:

$$L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & -4 & 0 & 1 \end{bmatrix}$$

Checking this:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 5 & 5 \\ 0 & 4 & 6 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 4 \end{bmatrix}$$

Question: What should our last matrix be? **Answer:**

$$L_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

Checking this:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

This matrix is exactly what we sought: it is upper triangular. So we can say that

$$U = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Which means that we have obtained the following equation:

$$L_3 L_2 L_1 A = U$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Note though that we're not quite done. We said we really wanted to factor A; we want two matrices which, when multiplied together, get A. To find this, we observe, as in the problem set, that $A = LU$ where $L = L_1^{-1}L_2^{-1}L_3^{-1}$. So to find the L that satisfies $A = LU$, we now have to find the inverses of the matrices L_k. This is a long process for 4×4 matrices, which we'll cover in a future section.

4 Section 4: 2020 October 13

Topics:

- Quick chat about grades in a PhD program[1]
- A humane method for finding matrix inverses[2]

After this class, I expect every student to be able to:

- Follow the steps to take a 3×3 matrix determinant
- Take a matrix inverse much more easily

BACK TO TABLE OF CONTENTS

① Quick chat about grades in a PhD program:

We started this lesson by talking through what grades mean in a PhD program. Grades have no material relevance to nearly all of our typical professional goals, like passing our required classes, applying to funding sources, or applying to jobs at the end of the degree. I argue that grades could matter in exactly the following situations:

i) When you're taking a dual degree that has a minimum grade requirement
ii) If you're somehow at risk of falling below the Rackham 3.0 average threshold to be In Good Standing, which is not a concern for anyone who puts effort into classes in our department
iii) If you manage to find a grant or funding source that somehow cares about the grades you've gotten in your PhD program
iv) If you later want to drop out of the program and get a job or apply to another program using the grades you got in this program
v) It has happened that people who really don't try in methods classes, to the tune of skipping a large number of problem sets, and thereby get lower than a B, are later encouraged not to minor in methods

② A humane method for finding matrix inverses:

Let's motivate a new method for finding matrix inverses with an example.

Example: Suppose we wanted to find the inverse of the following 3×3 matrix A:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$$

Using the inverse formula, we have to do the following four steps:
i) Find the Determinant of A, $|A|$
ii) Take the Transpose of A, A^{-1}
iii) Find the Determinant of all of the 2×2 Minor Matrices of A^{-1}, and calculate the Adjugate Matrix by finding all of the Cofactors
iv) Divide the Adjugate Matrix by the determinant of A, $|A|$

$$A^{-1} = \frac{1}{|A|}(-1)^{i+j}|M_{ji}|$$

Let's follow this procedure now to find the inverse of this matrix.

i) Find $|A|$. By the Laplace expansion, $|A|$ is given by

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} = 1 \cdot \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} - 2 \cdot \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} + 3 \cdot \begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix}$$

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix} = (3 \cdot 2 - 1 \cdot 1) - 2 \cdot (2 \cdot 2 - 1 \cdot 3) + 3 \cdot (2 \cdot 1 - 3 \cdot 3)$$

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix} = 5 - 2 - 21$$

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix} = -18$$

ii) Next, we take the transpose of our (symmetric) matrix A:

$$A^T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$$

iii) Now, find the Determinant of all of the 2×2 Minor Matrices of A^{-1}, and calculate the Adjugate Matrix by finding all of the Cofactors:

$$\text{Adj}(A) = \begin{bmatrix} 1 \cdot \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} & -1 \cdot \begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} & 1 \cdot \begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix} \\ -1 \cdot \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} & 1 \cdot \begin{vmatrix} 1 & 3 \\ 3 & 2 \end{vmatrix} & -1 \cdot \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} \\ 1 \cdot \begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix} & -1 \cdot \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} & 1 \cdot \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} \end{bmatrix}$$

$$\text{Adj}(A) = \begin{bmatrix} (3 \cdot 2 - 1 \cdot 1) & -(2 \cdot 2 - 1 \cdot 3) & (2 \cdot 1 - 3 \cdot 3) \\ -(2 \cdot 2 - 3 \cdot 1) & (1 \cdot 2 - 3 \cdot 3) & -(1 \cdot 1 - 2 \cdot 3) \\ (2 \cdot 1 - 3 \cdot 3) & -(1 \cdot 1 - 3 \cdot 2) & (1 \cdot 3 - 2 \cdot 2) \end{bmatrix}$$

$$\text{Adj}(A) = \begin{bmatrix} 5 & -1 & -7 \\ -1 & -7 & 5 \\ -7 & 5 & -1 \end{bmatrix}$$

iv) Finally, divide the Adjugate Matrix by the determinant of A, $|A|$:

$$A^{-1} = \frac{1}{|A|}\text{Adj}(A)$$

$$A^{-1} = -\frac{1}{18}\begin{bmatrix} 5 & -1 & -7 \\ -1 & -7 & 5 \\ -7 & 5 & -1 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} -\frac{5}{18} & \frac{1}{18} & \frac{7}{18} \\ \frac{1}{18} & \frac{7}{18} & -\frac{5}{18} \\ \frac{7}{18} & -\frac{5}{18} & \frac{1}{18} \end{bmatrix}$$

Theorem 2. *Life is short.*

Proof. According to The Center for Disease Control, the average American life span is 78.7 years. That's only 157 semesters. □

Theorem 3. *Consider a $k \times k$ invertible matrix A, for any $k \in \mathbb{N}$:*

$$A = \begin{bmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{bmatrix}$$

Now consider the concatenation of A together with $I_{k\times k}$ in augmented matrix form, as:

$$\left[\begin{array}{ccc|ccc} a_{11} & \dots & a_{1k} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} & 0 & \cdots & 1 \end{array}\right]$$

If the left-hand side of the augmented matrix (containing A) is brought to reduced row-echelon form, then the right-hand side contains A^{-1}.

Example: So now let's return to the previous Example and see if Theorem 3 does indeed make our lives easier like I promised.

Theorem 3 says that first we want to staple $I_{3\times 3}$ onto A in the form of an augmented matrix, like this:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 & 1 & 0 \\ 3 & 1 & 2 & 0 & 0 & 1 \end{array}\right]$$

Next, we want to bring it to reduced row-echelon form. We know how to do that: Gaussian elimination! So let's give it a go.

$$\xrightarrow{R_2-2R_1\to R_2} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -5 & -2 & 1 & 0 \\ 3 & 1 & 2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3-3R_1\to R_3}\left[\begin{array}{ccc|ccc}1 & 2 & 3 & 1 & 0 & 0\\ 0 & -1 & -5 & -2 & 1 & 0\\ 0 & -5 & -7 & -3 & 0 & 1\end{array}\right]$$

$$\xrightarrow{R_3-5R_2\to R_3}\left[\begin{array}{ccc|ccc}1 & 2 & 3 & 1 & 0 & 0\\ 0 & -1 & -5 & -2 & 1 & 0\\ 0 & 0 & 18 & 7 & -5 & 1\end{array}\right]$$

$$\xrightarrow{\frac{1}{18}R_3\to R_3}\left[\begin{array}{ccc|ccc}1 & 2 & 3 & 1 & 0 & 0\\ 0 & -1 & -5 & -2 & 1 & 0\\ 0 & 0 & 1 & \frac{7}{18} & -\frac{5}{18} & \frac{1}{18}\end{array}\right]$$

$$\xrightarrow{5R_3+R_2\to R_2}\left[\begin{array}{ccc|ccc}1 & 2 & 3 & 1 & 0 & 0\\ 0 & -1 & 0 & -\frac{1}{18} & -\frac{7}{18} & \frac{5}{18}\\ 0 & 0 & 1 & \frac{7}{18} & -\frac{5}{18} & \frac{1}{18}\end{array}\right]$$

$$\xrightarrow{2R_2+R_1\to R_1}\left[\begin{array}{ccc|ccc}1 & 0 & 3 & \frac{16}{18} & -\frac{14}{18} & \frac{10}{18}\\ 0 & -1 & 0 & -\frac{1}{18} & -\frac{7}{18} & \frac{5}{18}\\ 0 & 0 & 1 & \frac{7}{18} & -\frac{5}{18} & \frac{1}{18}\end{array}\right]$$

$$\xrightarrow{R_1-3R_3\to R_1}\left[\begin{array}{ccc|ccc}1 & 0 & 0 & -\frac{5}{18} & \frac{1}{18} & \frac{7}{18}\\ 0 & -1 & 0 & -\frac{1}{18} & -\frac{7}{18} & \frac{5}{18}\\ 0 & 0 & 1 & \frac{7}{18} & -\frac{5}{18} & \frac{1}{18}\end{array}\right]$$

$$\xrightarrow{R_1 - 3R_3 \to R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{5}{18} & \frac{1}{18} & \frac{7}{18} \\ 0 & 1 & 0 & \frac{1}{18} & \frac{7}{18} & -\frac{5}{18} \\ 0 & 0 & 1 & \frac{7}{18} & -\frac{5}{18} & \frac{1}{18} \end{array}\right]$$

So, just as with our previous method, we found

$$A^{-1} = \begin{bmatrix} -\frac{5}{18} & \frac{1}{18} & \frac{7}{18} \\ \frac{1}{18} & \frac{7}{18} & -\frac{5}{18} \\ \frac{7}{18} & -\frac{5}{18} & \frac{1}{18} \end{bmatrix}$$

5 Section 5: 2020 October 20

Topics:

- A return to the friendly LU factorization[1]
- Eigenvalues, minors, and definiteness[2]

After this class, I expect every student to be able to:

- Follow the steps to take a 4×4 matrix determinant
- Apply either the eigenvalue or leading submatrix method to check definitness

BACK TO TABLE OF CONTENTS

① **A return to the friendly LU factorization**:

Let's go back to the LU factorisation we were dealing with in a previous section, and let's check the result using the matrix inversion trick from last section. If you recall, we had a problem: we wanted to take the inverse of a product of 3 matrices, each of which was 4×4. That is, we had $L_3L_2L_1A = U$, and we wanted to find A which requires left-multiplying both sides by $L_1^{-1}L_2^{-1}L_3^{-1}$. The setup looked like this:

$$L_3L_2L_1A = U$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Sadly, as we have now seen, finding the full inverse requires finding the value of the determinant deleting **each combination of rows and columns**, from deleting the row 1 and column 1, row 1 and column 2, ..., all the way up to deleting row 4 and column 4. So the result looks like this:

$$A^{-1} = \frac{1}{|A|}(-1)^{i+j}|M_{ji}|$$

$$A^{-1} = \frac{1}{|A|} \begin{bmatrix} |M_{11}| & -|M_{21}| & |M_{31}| & -|M_{41}| \\ -|M_{12}| & |M_{22}| & -|M_{32}| & |M_{42}| \\ |M_{13}| & -|M_{23}| & |M_{33}| & -|M_{43}| \\ -|M_{14}| & |M_{24}| & -|M_{34}| & |M_{44}| \end{bmatrix}$$

Where $|M_{ij}|$ is the determinant of the submatrix obtained by taking A and removing row i and column j. Notice that the $+$ and $-$ signs alternate in the following nice pattern:

$$\begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix}$$

Let's just get a taste of the problem by making sure that L_1 is even invertible. So we want to find the determinant of this 4×4 matrix:

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{vmatrix} = L(1,1) \cdot L_1(-1,-1) - L(1,2) \cdot L_1(-1,-2) + L(1,3) \cdot L_1(-1,-3) - L(1,4) \cdot L_1(-1,-4)$$

After witnessing this brutal reality, I recommend that you take a moment, breathe, look out a window, relax, and only then continue.

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} - 0 \cdot \begin{vmatrix} -2 & 0 & 0 \\ -4 & 1 & 0 \\ -3 & 0 & 1 \end{vmatrix} + 0 \cdot \begin{vmatrix} -2 & 1 & 0 \\ -4 & 0 & 0 \\ -3 & 0 & 1 \end{vmatrix} - 0 \cdot \begin{vmatrix} -2 & 1 & 0 \\ -4 & 0 & 1 \\ -3 & 0 & 0 \end{vmatrix}$$

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} - 0 \cdot \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} + 0 \cdot \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix}$$

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{vmatrix} = 1 \cdot \frac{1}{1 \cdot 1 - 0 \cdot 0}$$

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{vmatrix} = 1$$

That could have been worse! So at least it is invertible. Really we should use this determinant to find the inverse, and then make sure that both of the other lower triangular matrices are also nonsingular! But, recalling Theorem 2, there has to be a better way than this. Thankfully, there is!

Exercise: Use our new trick to invert this matrix by means of Gauss-Jordan elimination.

Alternative approach: With Theorem 2 in mind, let's step back and see if there's any other simpler way to do this than to apply the full horrendous algorithm. And indeed, if we squint at the matrix L_1 for a while, we're likely to notice the following pattern. We want to find a matrix L_1^{-1} satisfying

$$L_1^{-1} L_1 = I_{4\times 4}$$

So all we need to do is eliminate the 2, 3, and 4 out of the leftmost column of L_1 without touching any of the other rows or columns, and we'll reduce it down to the identity matrix. Following this intuition leads us to the matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 4 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This turns out to be a general claim about this sort of matrix: when we have a lower triangular matrix with nonzero entries off the diagonal in just one column, the inverse of that matrix can be obtained by just flipping the minus signs on each element below the diagonal. If you like, go back and check that the formula we wrote out for an inverse will get us the result that we just claimed.

To finish the example, let's go back and find the inverse of both L_2 and L_3. These are similarly given by

$$L_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 4 & 0 & 1 \end{bmatrix}$$

And

$$L_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Finally, their product L is as clean as can be: we just mush together all of the nonzero entries in all of the matrices L_k:

$$\begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 4 & 3 & 1 & 0 \\ 3 & 4 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

(2) **Eigenvalues, minors, and definiteness**:

Let's now take a look at eigenvalues, minors, and definiteness.

Theorem 4. *A symmetric matrix is called each of the following if all of its eigenvalues* λ *satisfy:*

Positive definite iff $\lambda > 0$
Positive semi-definite iff $\lambda \geq 0$
Negative definite iff $\lambda < 0$
Negative semi-definite iff $\lambda \leq 0$

Let's check this out on some matrices of increasing size, starting with this 2×2 matrix:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

Then the characteristic equation (the equation we set equal to zero in order to solve for the eigenvalues) is given by:

$$0 = \det(A - I\lambda)$$

$$0 = \det\left(\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right)$$

$$0 = \det\left(\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right)$$

$$0 = \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix}$$

$$0 = (1-\lambda)(1-\lambda) - 2 \cdot 2$$

$$0 = 1 - 2\lambda + \lambda^2 - 4$$

$$0 = \lambda^2 - 2\lambda - 3$$

$$0 = \lambda^2 - 2\lambda - 3$$

$$0 = (\lambda - 3)(\lambda + 1)$$

So the eigenvalues of this matrix are $\lambda_1 = 3$ and $\lambda_2 = -1$. By the theorem, it is indefinite. Let's confirm that using the theorem that Iain introduced in lecture, where we check definiteness using the sign of the leading principle minor:

$$|L_1| = 1$$

$$|L_2| = 1 \cdot 1 - 2 \cdot 2$$

$$|L_2| = -3$$

Recall the theorem from lecture:

Theorem 5. *Let A be an $n \times n$ symmetric matrix. Then*

A is positive definite iff all its n leading principal minors are strictly positive

A is negative definite iff all its n leading principal minors alternate in sign such that $sign(|A_k|) = (-1)^k$

If A breaks those patterns in a way that isn't just a non-strict inequality, then it is indefinite

In this situation $|L_1| > 0$ but $|L_2| < 0$ is the opposite of the pattern that $\text{sign}(|A_k|) = (-1)^k$. So by this method too we see that the matrix is indefinite.

Which of our two methods for checking definiteness of the matrix is better? We discussed this briefly in section, but we didn't get as far as actually verifying why it's hard to use the characteristic equation to find the eigenvalues of a 3×3 matrix. So, with the caveat that the rest of this material was **not covered in section**, let's see what happens when we scale these procedures up a bit with the following simple-looking 3×3 matrix:

$$\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

First let's use the eigenvalue method. So we need to solve the caracteristic equation:

$$0 = \det(A - I\lambda)$$

$$0 = \det\left(\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 3 \\ 2 & 3 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \right)$$

$$0 = \begin{vmatrix} 1-\lambda & 2 & 2 \\ 2 & 1-\lambda & 3 \\ 2 & 3 & 1-\lambda \end{vmatrix}$$

$$0 = (1-\lambda) \begin{vmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{vmatrix} - 2 \begin{vmatrix} 2 & 3 \\ 2 & 1-\lambda \end{vmatrix} + 2 \begin{vmatrix} 2 & 1-\lambda \\ 2 & 3 \end{vmatrix}$$

$$0 = (1-\lambda)((1-\lambda)(1-\lambda) - 3 \cdot 3) - 2(2 \cdot (1-\lambda) - 2 \cdot 3) + 2(2 \cdot 3 - (1-\lambda) \cdot 2)$$

$$0 = (1-\lambda)(1 - 2\lambda + \lambda^2 - 9) - 2(2 - 2\lambda - 6) + 2(6 - 2 + 2\lambda)$$

$$0 = (1-\lambda)(\lambda^2 - 2\lambda - 8) - 2(-2\lambda - 4) + 2(4 + 2\lambda)$$

$$0 = \lambda^2 - 2\lambda - 8 - \lambda^3 + 2\lambda^2 + 8\lambda + 4\lambda + 8 + 8 + 4\lambda$$

$$0 = -\lambda^3 + 3\lambda^2 + 14\lambda + 8$$

The roots of this equation are something terribly messy: they turn out to be $\lambda_1 = -2, \lambda_2 = \frac{1}{2}(5 - \sqrt{41}), \lambda_3 = \frac{1}{2}(5 + \sqrt{41})$. With two negative and one positive eigenvalues, this is an indefinite matrix. In contrast to this mess, you can see that the signs of the leading principle minors won't be so bad:

$$|L_1| = 1$$

$$|L_2| = \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix}$$

$$|L_2| = -3$$

Already we know that this has to be an indefinite matrix. Notice a little trick here: if we already see that the patterns are broken, we know immediately it's an indefinite matrix, without even having to calculate the determinant of the full 3×3 matrix.

In my experience the leading principal minors method is usually faster and easier to compute by hand, *but* it requires you to remember more information to apply it correctly – the rules for using eigenvalues are much more straightforward (to my mind).

6 Section 6: 2020 October 27

Topics:

- Sets, mappings, and set sizes[1]
- Countability and uncountability[2]

After this class, I expect every student to be able to:

- define 1 to 1 and onto
- describe how to use 1 to 1 and onto correspondences to compare set sizes
- summarize the difference between "discrete" and "continuous" using the idea of sets and mappings

BACK TO TABLE OF CONTENTS

① **Sets and mappings**:

Sets are the fundamental building blocks of modern math.

You already know that numbers can be arranged into sets: $\mathbb{N}, \mathbb{R}, \mathbb{Z}, \mathbb{C}$, and so on. But a crucial observation is that many of the functions and procedures that you've studied in 598 and 599 can be written as mappings from a set of numbers to a set of numbers. Let's look at some examples.

Example: The function n^2 for $n \in \mathbb{N}$ maps as follows:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ \vdots \end{pmatrix} \xrightarrow{\quad n^2 \quad} \begin{pmatrix} 1 \\ 4 \\ 9 \\ \vdots \end{pmatrix}$$

This map will only ever produce natural numbers, so it maps from the set of natural numbers to the set of natural numbers. We write this fact as $n^2 : \mathbb{N} \to \mathbb{N}$.

Definition: A correspondence c from set S to T is called **1 to 1** or **injective** if c never maps multiple elements of S onto the same element of T.

Note: In class I drew the following examples visually using matrices and lines connecting the individual elements. Theoretically I know how to do that in LaTeX too, using the tikz package, but holy cow would it ever be a colossal pain. If you really want to see it, email me and I'll see what I can do.

Example: Say $S = \{1, 2, 3\}$, $T = \{A, B, C, D\}$, and we have
$c(1) = A$
$c(2) = B$
$c(3) = C$
This is a 1-1 function.

Example: Say $S = \{1, 2, 3\}$, $T = \{A, B, C, D\}$, and we have
$c(1) = A$
$c(2) = B$
$c(3) = B$
This is not a 1-1 function.

Definition: A correspondence c from set S to set T is called **onto** or **surjective** if every element in T is mapped to by at least one element in S.

Example: Say $S = \{1, 2, 3, 4\}$, $T = \{A, B, C\}$, and we have
$c(1) = A$
$c(2) = B$

$c(3) = C$
$c(4) = C$
This is an onto function.

Question: Is it 1-1? **Answer:** No, because 3 and 4 both output C.

Example: Say $S = \{1, 2, 3\}$, $T = \{A, B, C, D\}$, and we have
$c(1) = A$
$c(2) = B$
$c(3) = C$
This is not an onto function.

Question: Is it 1-1? **Answer:** Yes.

Definition: A set relation is called **bijective** (or, simply, **1 to 1 and onto**) if it is injective (1 to 1) and surjective (onto).

Question: Were any of the examples so far 1-1 and onto? **Answer:** No.

Example: Say $S = \{1, 2, 3, 4\}$, $T = \{A, B, C, D\}$, and we have
$c(1) = A$
$c(2) = B$
$c(3) = C$
$c(4) = D$
This is a 1 to 1 and onto function.

Example: We said that this is useful because sets and their relationships are the fundamental building blocks of math. So let's take this back to more familiar terrain for a moment just to see the connection. Consider the function n^2 for $n \in \mathbb{N}$. Is this map 1 to 1? Is it onto? **Answer** It is 1 to 1 because every natural number has a square. It is not onto because not every natural number has a square root within the set of natural numbers. For example, we can never square a natural number to obtain $\sqrt{5}$. Notice also that the square root of positive integers is not 1 to 1, because for example $\sqrt{25} = 5$ and $\sqrt{25} = -5$, but $5 \neq -5$, there is both a positive and a negative root.

Example: Is $n + 1$ 1 to 1 and onto for $n \in \mathbb{N}$? **Answer:** It is 1 to 1, but it is not onto, because it misses the smallest natural number.

Question: Can anyone suggest a mapping $f : \mathbb{N} \to \mathbb{N}$ that is 1 to 1 and onto? **Hint:** Think simple. **Answer:** $f(n) = 1 \cdot n$ suffices!

Question: How about between $\mathbb{R}^+$ and $\mathbb{R}^-$?

Exercise: Take any two sets S and T of different cardinalities (so, containing different numbers of elements). Make a mapping that is 1 to 1 and onto between these two sets.

Answer: It is not possible to do so!

Theorem 6. *Sets S and T satisfy $|S| = |T|$ iff there is a mapping $f : S \to T$ which is both 1 to 1 and onto.*

In many presentations (Rosenthal et al., 2014: p. 87), this is actually not a theorem but rather a definition of cardinality. To see why these two ideas are the same, let's try to see why your attempts to make a 1 to 1 and onto function between two sets of different sizes didn't work.

Suppose that $|S| > |T|$. Then in order to make the correspondence 1 to 1, you are going to have to re-use elements in T. This means that if the correspondence is 1 to 1, it cannot be onto, so it can never be both 1 to 1 and onto. Now suppose that $|S| < |T|$. In order to make the correspondence onto, at least one element in S is going to have to map onto multiple elements of T, which means that it is not a successful mapping; it fails to tell us what to do to an element of S to get it into T. This means that the only case in which the correspondence can be both 1 to 1 and onto is the case in which $|S| = |T|$.

(2) **Countability and uncountability**:

↪ We have talked about "countable" and "uncountable" sets. The difference between these two, and the connection between them and "discrete" or "continuous" variables, is one of the easiest, most powerful, and most beautiful proofs in modern math.

Remark: Discrete variables are defined on some (maybe improper) subset of the set of all integers ($\mathbb{Z}$, and recalling that $\mathbb{N} \subset \mathbb{Z}$), and continuous variables are defined on a larger number set like the set of real numbers ($\mathbb{R}$). We call the numbers that a discrete variable is defined on "countable" (simply because natural numbers are the numbers that humans count with), and the numbers that a continuous variable is defined on "uncountable". This naming convention is explained by the following theorem:

Theorem 7. *It is not possible to associate every natural number with a real number, without leaving any real numbers out.*

Note: This theorem says that there are more real numbers than there are integers or natural numbers, even though these sets of numbers are infinite.

Proof. First construct a countably infinite list of real numbers:[3]

[3]This step in the proof used to be very controversial, and a lot of historical effort went into showing that such a list can be constructed. The rigorous proof for the existence of an arbitrary list of distinct reals in decimal form turns out to be extremely difficult. If you're interested in this, Kline (1972) has a really wonderful discussion.

$$\begin{array}{ccccccccc}
1 & . & 0 & 0 & 0 & 0 & 0 & \cdots \\
1 & . & 0 & 0 & 0 & 1 & 0 & \cdots \\
3 & . & 1 & 4 & 1 & 5 & 9 & \cdots \\
5 & . & 9 & 8 & 5 & 9 & 9 & \cdots \\
2 & . & 1 & 0 & 1 & 0 & 1 & \cdots \\
2 & . & 7 & 8 & 1 & 8 & 2 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

When I call this list "countably infinite", I mean that we can associate every number in it with a natural number, which means that we can count the numbers in the list. I can refer to the first number in the list, the second number in the list, the hundredth number in the list, and so on. So, this list is no larger than the set of natural numbers.

Now, consider the ith digit of each of number i in the list, so the 1st digit of the 1st number, the 2nd digit of the second number, and so on:

$$\begin{array}{ccccccccc}
\textcircled{1} & . & 0 & 0 & 0 & 0 & 0 & \cdots \\
1 & . & \textcircled{0} & 0 & 0 & 1 & 0 & \cdots \\
3 & . & 1 & \textcircled{4} & 1 & 5 & 9 & \cdots \\
5 & . & 9 & 8 & \textcircled{5} & 9 & 9 & \cdots \\
2 & . & 1 & 0 & 1 & \textcircled{0} & 1 & \cdots \\
2 & . & 7 & 8 & 1 & 8 & \textcircled{2} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

Let's consider the number, call it r, that is formed by concatenating each of those digits together:

$$r = 1.04502 \cdots$$

Now let's modify every digit of this number, say by adding 1 to every non-9 digit, and sending every 9 to 0. So we obtain a new real number, call it r':

$$r' = 2.15613 \cdots$$

r' is real, but it is not in our countably infinite list, because its first digit differs from the first digit of the first number, its second digit differs from the second digit of the second number, and so on.

Therefore, given any countably infinite list of real numbers, this procedure can always produce a real number which was not in our countably infinite list. No matter how many countably infinite numbers we list, we can always produce a number which is not in that list. So any mapping from natural numbers to real numbers will never cover all of the real numbers.

□

So, by Theorem 6, we've just learned something interesting: there is no such thing as a bijective mapping $f : \mathbb{N} \rightarrow \mathbb{R}$. Is this true for all the really big number sets that we care about?

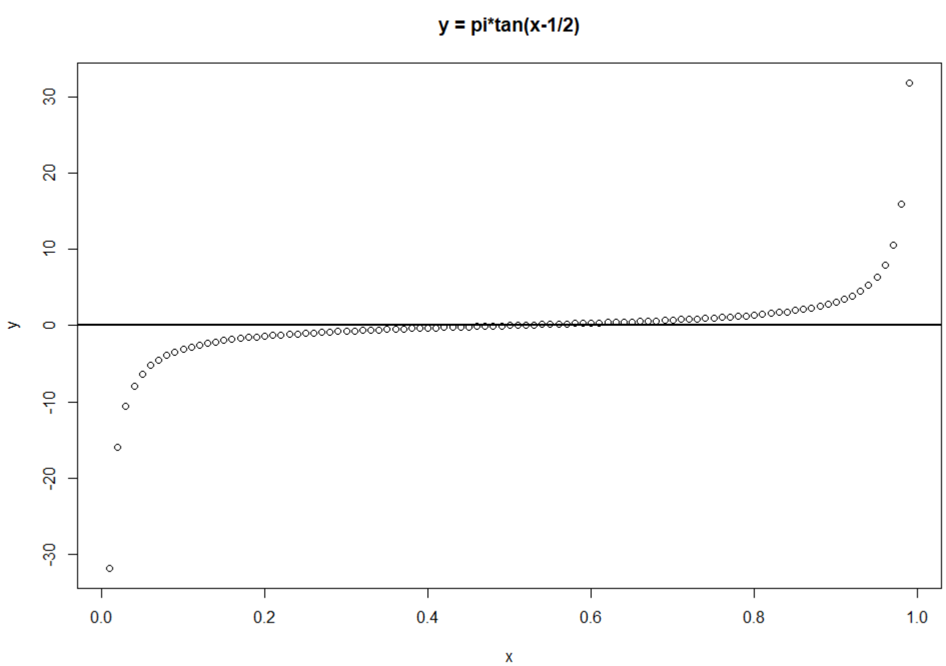
Interestingly enough, not at all! There are tons of extremely bizarre mappings that we can find. Since it's almost Halloween, let's see some super spooky mappings.

Example: The cardinality of the set of even integers is the same as the cardinality of the set of integers. We simply use the bijective mapping $f(z) = 2z \ \ \forall z \in \mathbb{Z}$.

Example: We can show that $|\mathbb{Z}| = |\mathbb{N}|$ by mapping every even natural number to the next consecutive positive integer and every odd natural number to the next consecutive negative integer.

Example: The following mapping shows that $|(0;1)| = |\mathbb{R}|$:

```
BY = 0.01
x = seq(0+BY,1-BY,by=BY)
y = tan((x-1/2)*pi)
plot(x,y, main = "y = tan((x-1/2)*pi)")
abline(h=0,lwd=2)
```



Example: Consider the set of natural numbers $\mathbb{N}$ and the subset of reals in the interval $[0;1]$, and any 1 to 1 mapping $f : \mathbb{N} \to [0;1]$. Such a mapping cannot also be onto. To get some intuition for why, notice that there are irrational real numbers in the interval $[0;1]$, such as $\frac{1}{2}\sqrt{2}$. Probably the most straightforward way to create f is to define ways of diving naturals by each other, but that can never map onto the irrational reals in $[0;1]$.

7 Section 7: 2020 November 3

Topics:

- Using mountain climbing to interpret optimisation[1]

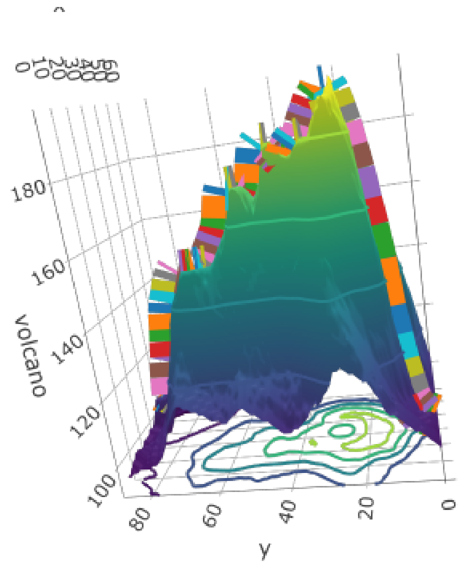
After this class, I expect every student to be able to:

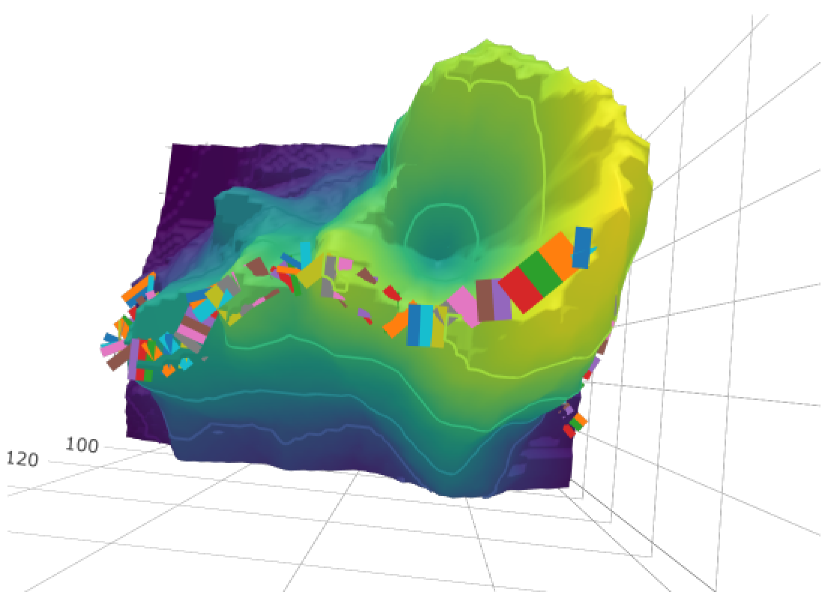
- Associate a clear picture with partial and directional derivatives
- Be able to describe partial derivatives with reference to a 3D surface

BACK TO TABLE OF CONTENTS

① **Using mountain climbing to interpret partial derivatives**:

This section was a thorough tour of iterative optimisation methods using a 3D plot of the volcano matrix in R. We followed a mountain climber up the volcano using a series of different rulesets to attach some strong visuals to the mathematical ideas from lecture. Some still examples of the 3D plot follow, and the code is available on Canvas.[4]





[4]Or on request.

8 Section 8: 2020 November 10

Topics:

- Sign of the Lagrangian[1]
- Lagrangians in statistics[2]

After this class, I expect every student to be able to:

- construct the Lagrangian for a constrained optimization problem by either adding or subtracting the Lagrangian multiplier
- remember that constrained optimisation is definitely not just for game theorists!

BACK TO TABLE OF CONTENTS

① **Sign of the Lagrangian**:

Note: When seeking $\max\limits_{x_1,\ldots,x_j} f(x_1,\ldots,x_j)$ under the k equality constraints $g_1(x_1,\ldots,x_j),\ldots,g_k(x_1,\ldots,x_j)$, the following two definitions of the Lagrangian are exactly equivalent:

$$\mathcal{L} = f(x_1,\ldots,x_j) + \lambda_1 g_1(x_1,\ldots,x_j) + \ldots + \lambda_k g_k(x_1,\ldots,x_j)$$

$$\mathcal{L} = f(x_1,\ldots,x_j) - \lambda_1 g_1(x_1,\ldots,x_j) - \ldots - \lambda_k g_k(x_1,\ldots,x_j)$$

To make sure that notation isn't getting in the way, take the example of two variables x and y and one equality constraint $g(x,y)$ on the function $f(x,y)$ to be maximized. Then the Lagrangian is either of the following:

$$\mathcal{L} = f(x,y) + \lambda g(x,y)$$

$$\mathcal{L} = f(x,y) - \lambda g(x,y)$$

Example: Let's try it both ways with an optimization problem. Let's say that we want to find the point on the line $x + y = 1$ which is closest to the point (2,3).[5]

To set up the constrained optimization problem, we need two pieces. First, we need the function $f(x,y)$ to be maximized. This is the distance between any given coordinate (x,y) and the point $(2,3)$. We've talked a lot about the default distance equation in 2-dimensional space: the Euclidean 2-norm. So that's

$$f(x,y) = \sqrt{(x-2)^2 + (y-3)^2}$$

But we have the opportunity to simplify $f(x,y)$ here – and this is always a desirable thing to do, as we learned from Theorem 2. With a small amount of algebra we can see that $\sqrt{a} > \sqrt{b} \iff a > b$ for $0 < a, b \in \mathbb{R}$. So we have the very nice opportunity to simply drop the square root and take our objective function to be:

$$f(x,y) = (x-2)^2 + (y-3)^2$$

Note: This is an extremely fundamental trick in scientific computing, because square roots have to be found by an extremely slow iterative algorithm. If you are comparing the sizes of quantities in a computer program and the program is running very slowly, ask yourself:

[5] I made this specific example up completely, but lots of excellent examples along these lines (although not exactly focusing on the sign of the Lagrangian) can be found in the standard real analysis text by Edwards Jr. (1973).

are there any square roots here that I can drop without changing the answer?

And we were given the constraint that the point must be on the line $x + y = 1$. The claim above says that we can choose either of the following functions to be $g(x, y)$:

$$g(x, y) = x + y - 1$$

$$g(x, y) = 1 - x - y$$

So let's try it with each constraint, starting with the first one (**Note:** If you look at the Lagrangian you will see that trying it with each constraint is identical to trying it with $\pm\lambda$). The full Lagrangian is

$$\mathcal{L} = (x - 2)^2 + (y - 3)^2 + \lambda(x + y - 1)$$

$$\mathcal{L} = x^2 - 4x + 4 + y^2 - 6y + 9 + \lambda x + \lambda y - \lambda$$

$$\mathcal{L} = x^2 + y^2 - 4x - 6y + 13 + \lambda x + \lambda y - \lambda$$

Then we have the following partial derivatives:

$$\begin{bmatrix} \mathcal{L}_x \\ \mathcal{L}_y \\ \mathcal{L}_\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} 2(x - 2) + \lambda \\ 2(y - 3) + \lambda \\ x + y - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

From the last equation, we have

$$x + y - 1 = 0$$

$$x + y = 1$$

And the first two equations say

$$2(x - 2) + \lambda = 0$$

$$x = -\frac{\lambda}{2} + 2$$

And symmetrically

$$y = -\frac{\lambda}{2} + 3$$

Plugging these latter two identities into the first gives

$$1 = -\frac{\lambda}{2} + 2 - \frac{\lambda}{2} + 3$$

$$1 = -2\frac{\lambda}{2} + 5$$

$$-4 = -\lambda$$

$$\lambda = 4$$

Using this value of λ to find x and y then,

$$x = -\frac{4}{2} + 2$$

$$x = 0$$

And

$$y = -\frac{\lambda}{2} + 3$$

$$y = 1$$

This was the unique critical point.[6] So was this a maximum or a minimum? To see this we have to construct the Bordered Hessian:

$$B = \begin{bmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{bmatrix}$$

[6]We do not need to check the constraint qualification to know that this is the unique critical point, because as we saw in lecture, whenever the constraint function is linear the constraint qualification is satisfied and the first order conditions of the Lagrangian will uncover every critical point.

$$B = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

So that

$$|B| = 0(2 \cdot 2 - 0 \cdot 0) - 1(1 \cdot 2 - 1 \cdot 0) + 1(1 \cdot 0 - 1 \cdot 2)$$

$$|B| = -2 - 2$$

$$|B| < 0$$

In lecture we saw that, with $k = 1$ equality constraints, then $(-1)^1|B| > 0 \implies$ minimum. Since $|B| < 0 \implies -1|B| > 0$, therefore the critical point is a minimum.

So that completes this first way of solving the problem. Now what if we had rearranged the constraint differently: what if we had instead picked the second constraint, so that the Lagrangian is

$$\mathcal{L} = (x-2)^2 + (y-3)^2 + \lambda(1 - x - y)$$

So what happens to the partial derivatives? Now we have

$$\begin{bmatrix} \mathcal{L}_x \\ \mathcal{L}_y \\ \mathcal{L}_\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 2(x-2) - \lambda \\ 2(y-3) - \lambda \\ 1 - x - y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So this time the last equation gives

$$1 - x - y = 0$$

$$x + y = 1$$

And the first two equations say

$$2(x-2) - \lambda = 0$$

$$x = \frac{\lambda}{2} + 2$$

And symmetrically

$$y = \frac{\lambda}{2} + 3$$

Combining these three expressions,

$$1 = \frac{\lambda}{2} + 2 + \frac{\lambda}{2} + 3$$

$$1 = 2\frac{\lambda}{2} + 5$$

$$\lambda = -4$$

Then, plugging this value into the equations for x and y,

$$x = \frac{\lambda}{2} + 2$$

$$x = -\frac{4}{2} + 2$$

$$x = 0$$

and similarly

$$y = \frac{\lambda}{2} + 3$$

$$y = -\frac{4}{2} + 3$$

$$y = 1$$

Question: How exactly did this work out? **Answer:** When we take all of the partial derivatives, no matter how we chose to rearrange the constraint, the constraint itself will always be the same. The difference that will be enforced is whether the partial derivatives with respect to the choice variables results in adding or subtracting the term(s) that include(s) λ. In this case, if λ was positive, the partial derivatives would work out so that we are subtracting it. If λ is negative, the partial derivatives work out so that we are adding it. The critical point is the same regardless.

Next let's just make sure that it remains the minimizer. To do this we seek the bordered Hessian of the function $\mathcal{L} = (x-2)^2 + (y-3)^2 + \lambda(1-x-y)$, which is given by:

$$B = \begin{bmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

so that the determinant is now given by

$$|B| = 0 \cdot \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} - (-1) \cdot \begin{vmatrix} -1 & 0 \\ -1 & 2 \end{vmatrix} + (-1) \cdot \begin{vmatrix} -1 & 2 \\ -1 & 0 \end{vmatrix}$$

$$|B| = 0(2 \cdot 2 - 0 \cdot 0) - (-1)(-1 \cdot 2 - (-1) \cdot 0) + (-1)((-1) \cdot 0 - (-1) \cdot 2)$$

$$|B| = -2 - 2$$

$$|B| < 0$$

So we have now shown that for this example, it doesn't matter how you rearrange the constraint equation $g(x, y)$, and it also doesn't matter whether you add or subtract $\lambda g(x, y)$ when you form the Lagrangian. This is true in general, not just for this example.

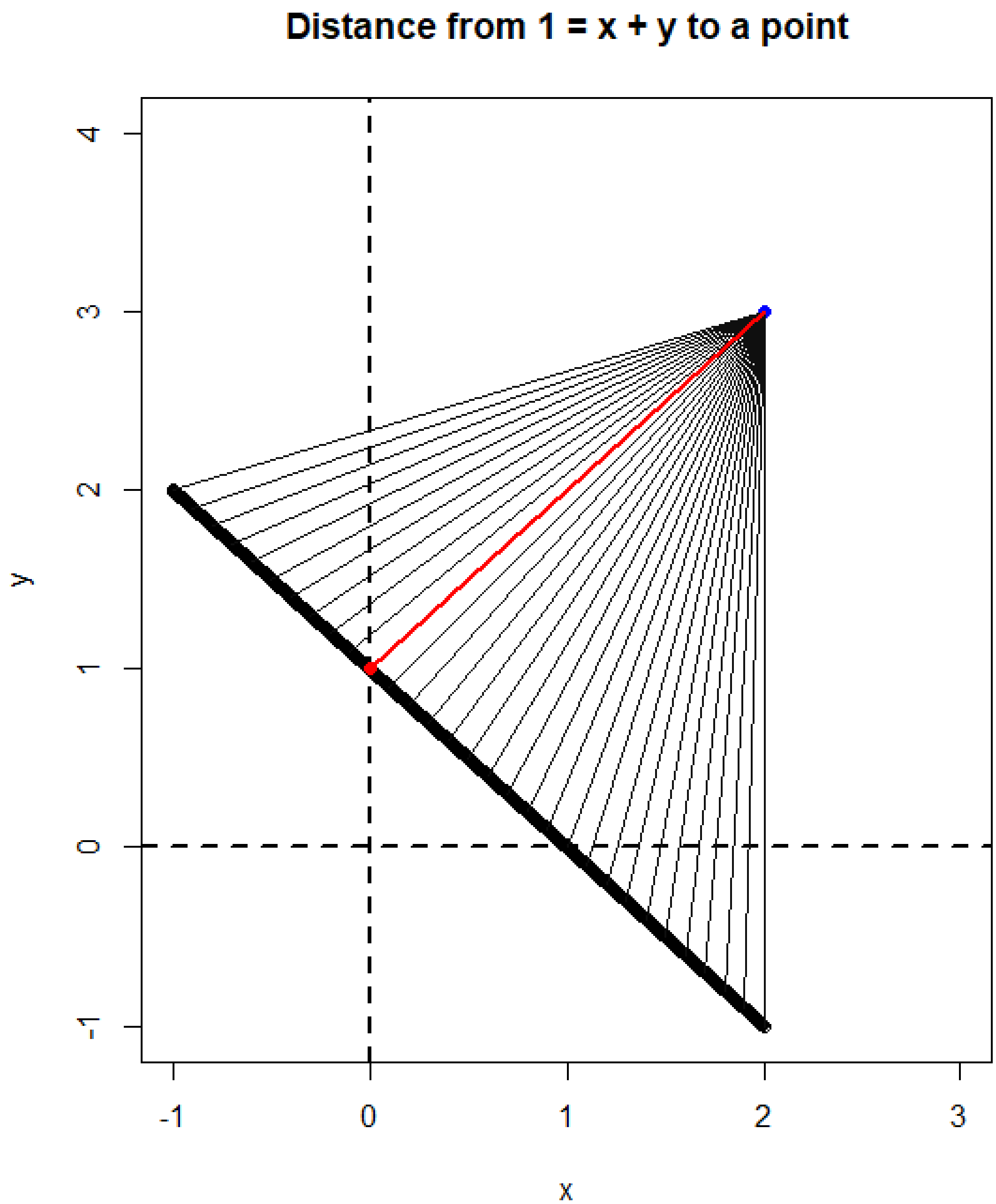
Question: Now, how could we have identified a maximizer of this function? **Answer:** There is no maximizer; we can set x and y to be arbitrarily far away along this line. If the constraint does not define a compact region, then we may have only one critical point, a maximum or a minimum. But if the constraint defined a compact region (for example, if I had added the second constraint that $x, y > 0$, or if we were on a circle instead of a line), then we would necessarily have both at least one maximum and at least one minimum.

The following R syntax and plot shows what exactly is going on here:

```
#Global vars
LINES <- TRUE    #Whether or not to draw lines to the point
PLOT_EVERY_N <- 100    #How many lines to plot

#Location of the point to find the distance to
```

```
PT_X <- 2
PT_Y <- 3

#x-value range to consider
X_MIN <- -1
X_MAX <- 2

#Construct the set of x points and y points according to 1 = x + y
xPoints <- seq(X_MIN,X_MAX,by=0.001)
yPoints <- 1 - xPoints

#Draw the plot without lines
plot(xPoints,yPoints, xlim=c(X_MIN, max(X_MAX, PT_X+1)),
        ylim=c(min(yPoints), max(1, PT_Y+1)),
        xlab="x",ylab="y",main="Distance from 1 = x + y to a point")
points(PT_X, PT_Y, col='blue', pch=19)
#Add dashed axes to plot
abline(v=0, lwd=2, lty=2)
abline(h=0, lwd=2, lty=2)

#Calculate a list of all of the distances at each line point
dist <- (PT_X - xPoints)**2 + (PT_Y - yPoints)**2

#If the LINES switch is on, then draw lines representing those distances
if (LINES) {
  for (i in seq(length(dist))) {
    if (!((i - 1) %% PLOT_EVERY_N)) {
      lines(c(xPoints[i],PT_X),c(yPoints[i],PT_Y), col = "grey6", lwd=1.5)
    }
  }
}

#Find the point corresponding to the minimum of all distances
minX <- xPoints[which(dist == min(dist))]
minY <- yPoints[which(dist == min(dist))]

#Colour the minimum line and minimum point red so they stand out
if (LINES){
  lines(c(minX,PT_X),c(minY,PT_Y), col = "red", lwd=2)
  points(minX,minY,col="red",pch=19)
}

#Print the resultant value
print(
  paste0(
```

```
    "The point on the line y + x = 1 which is closes to (", PT_X,",",PT_Y,
    ") is ","(",minX,",",minY,")"
  )
)
```


Distance from 1 = x + y to a point
y
x
4
3
2
1
0
-1
-1
0
1
2
3

(2) **Lagrangians in statistics**:

Remark: We have focused very heavily on applications of Lagrangians in formal modeling, because they're an absolutely inescapable tool that you need in order to get even some of the most basic results in game theory. But we can't let that obscure a complementary truth: that constrained optimisation is also essential for applied statisticians. I want to just quickly review some stats tools that explain why this is so. For reasons relating to my personal biases and background, I have mostly picked machine learning examples,[7] but rest assured that constrained optimisation turns up all over the rest of Bayesian statistics and in many hypothesis testing and inference situations.

Example: (Ridge regression) Without scooping POLSCI 699 by engaging in a detailed discussion of regression, this is a modification of the regression idea that is appropriate in some situations where regression fails to produce a single best fit line. In some cases where a regression might output multiple possible answers, ridge regression uses a constraint that reduces it to only one answer. Ridge regression is extremely important for political science (actually I think that political scientists use Ordinary Least Squares in many situations where they should use ridge regression, but let's talk about that only after you've finished 699), and the ridge regression equation is exactly a Lagrangian that looks just like the ones we've been studying.

Example: (Probabilities) We know that a probability p is bounded so that $0 \leq p \leq 1$ by definition. So often we want to maximize something involving probabilities under the constraints that $p \geq 0$ and $p \leq 1$. We can do this with arbitrarily many constraints ... just remember that of course we need more choice variables than constraints![8] A very common thing to maximize in Bayesian statistics is the likelihood that our hypothesis is the correct one, given the set of observed data.

Example: (Empirical risk minimization) Consider data which we divide into "training" and "test" data (the central trick in statistical classification), and we want to minimize the risk of misclassification. So we define something called the empirical risk: we classify the test data according to some procedure, and then we check the misclassification rate in the test data, and then we re-apply some modified version of the procedure to continually reduce the risk of misclassifying data points. However, this idea invites the very bad problem of overfitting: we will get a classification rule that is extremely well-suited to the particular data sample that we have, but might not generalize to other samples from the same population. One extremely common solution to this is to apply something called a "complexity penalty" to whatever procedure we come up with. So with observations θ, empirical risk $R(\theta)$, and the "procedure" (allow me to a bit a bit hand-wavy in what this procedure is exactly) $f(\theta)$, we introduce a complexity penalty $g(\theta)$ so that we have a constrained minimization problem of the form

[7] Relying heavily on the absolutely wonderful book Murphy (2012).

[8] A particularly nice discussion of this can be found in (Mathews and Walker, 1970: circa p. 332).

$$\min R(\theta, \lambda) = \min(f(\theta) + \lambda g(\theta))$$

This idea underlies probably a majority of the classification techniques and learning rules that you will encounter as a political scientist.

Example: (Support vector machines) This is a method for classifying points. We are trying to draw a separating line (called, to account for the case where we want to classify into more than 2 bins, a "separating hyperplane") to classify points into some number of categories. More than one hyperplane will do this for pretty much any realistic dataset, so we also have the rule that we want the widest possible margin between the hyperplane and the nearest data points, to give us a better chance of correctly classifying out-of-sample points which might be close to the boundary we're drawing between the two categories. This is a classic constrained optimisation problem, and one which sort of flips the examples we've seen on their heads: the function to be maximized is the margin size, and the constraint is that we have to maximize that while minimizing the misclassification rate. Conversely, the function to be maximized could be the classification rate, while also maximizing the margin size. So this is a fun example where our constraint is actually to maximize a whole other quantity! It collapses to a simple Lagrangian if you just say that the constraint is that the margin has to be bigger than a certain value.

9 Section 9: 2020 November 17

Topics:

- Origins of utilities and optimisation[1]
- Optimisation in games[2]

After this class, I expect every student to be able to:

- explain how utilities, optimisation, and strategic games connect
- describe the idea of spatial utilities and write down a simple functional form for them

BACK TO TABLE OF CONTENTS

(1) **Origins of utilities and optimisation**:

Note: This part of 598 is about taking all of the tools that we developed at the start of the course relating to derivatives, and showing how they can be used to optimise all sorts of different functions. One extremely important result is that strictly concave functions can be easily optimized, since they are guaranteed to have a unique maximum. These are the sorts of functions where our mountain climber is completely guaranteed to reach the top of the mountain just by following the steepest upwards-pointing derivative.[9]

Definition 1. *Spatial utilities are a method of assigning utilities to an agent who faces a decision problem by placing them at some point in a space, also assigning every object of their choice to a point in that space, and defining a mapping between distance in space and utility of the agent such that closer objects are always preferred to farther objects.*

Note: This idea is really familiar in political science, thanks to Downs (1957).

Note: One common way of operationalizing this is to say that, if the decision-maker's position is v and they are calculating their utility for an object of a choice which is positioned at c, then the decision-maker's utility is simply given by

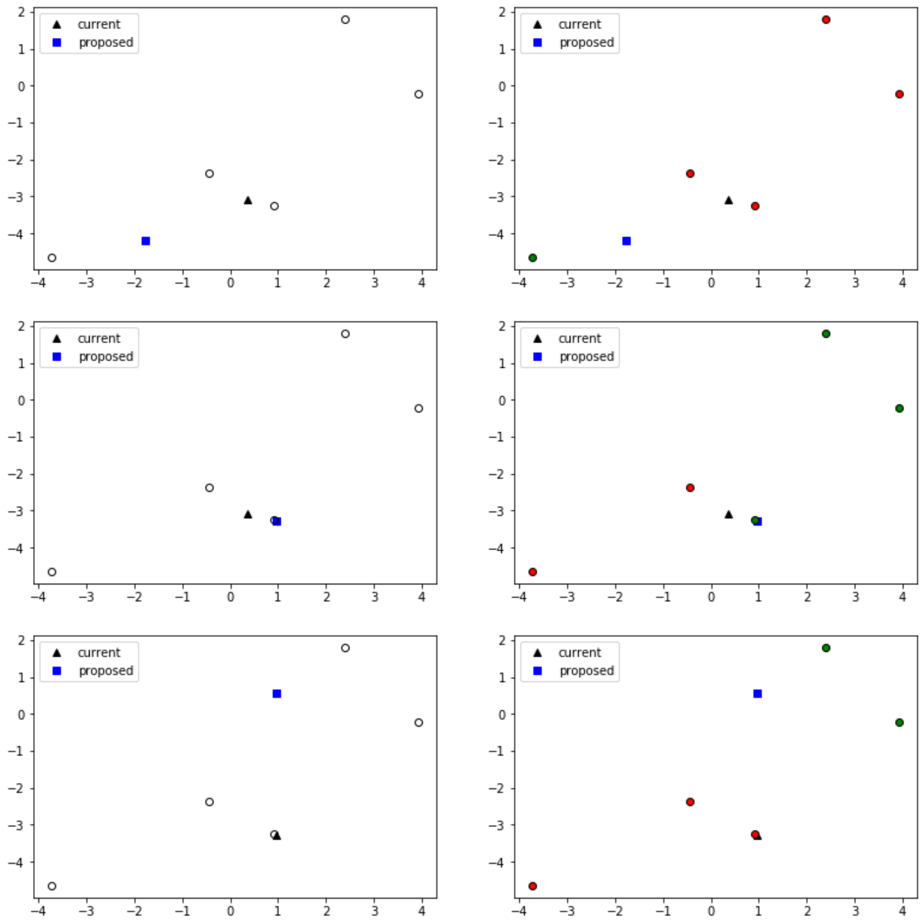
$$u(v, c) = -|v - c|$$

This is strictly increasing in the distance between v and c, with a maximum at 0 when $v = c$. If we had both an x and a y dimension, then we can use our favourite tool for deciding the distance between points in Euclidean space: the 2-norm. That looks like:

$$u(v, c) = -\sqrt{(x_v - x_c)^2 + (y_v - y_c)^2}$$

Example: Here are some images which demonstrate the decision problem of legislators voting between a current piece of legislation and a proposed amendment to that legislation.

[9] A related exposition to this one is in my 681 section notes from Winter 2019, available on my website.



It turns out that we can represent this as a problem of optimising a strictly concave function in 2 dimensions. Namely, that the spatial utility function of a voter positioned at (x_v, y_v) has a unique maximum at exactly their ideal point, (x_v, y_v). Consider a candidate at (x_c, y_c). We said that the voter's spatial utility obtained from voting for that candidate is:

$$u(v, c) = -\sqrt{(x_v - x_c)^2 + (y_v - y_c)^2}$$

$$u(v, c) = -\left((x_v - x_c)^2 + (y_v - y_c)^2\right)^{\frac{1}{2}}$$

First let's find the maximum on the x-dimension. Remember from Iain's discussion of concave utility functions that when we're trying to find a critical point on a function where we know that the unique critical point is just the unique maximum, we can simply take the derivative with respect to the variable of interest and set the function equal to 0:

$$\frac{\partial}{\partial x_c} u(v, c) = -\frac{1}{2}\left((x_v - x_c)^2 + (y_v - y_c)^2\right)^{-\frac{1}{2}} \cdot 2(x_v - x_c)(-1)$$

$$0 = \frac{x_v - x_c}{\sqrt{(x_v - x_c)^2 + (y_v - y_c)^2}}$$

Isolating x_c to figure out where the candidate should locate themselves, we get

$$0 = x_v - x_c$$

$$x_c = x_v$$

Similarly for the y-dimension,

$$\frac{\partial}{\partial y_c} u(v, c) = \frac{1}{2}\left((x_v - x_c)^2 + (y_v - y_c)^2\right)^{-\frac{1}{2}} \cdot 2(y_v - y_c)(-1)$$

$$0 = \frac{y_v - y_c}{\sqrt{(x_v - x_c)^2 + (y_v - y_c)^2}}$$

$$0 = y_v - y_c$$

$$y_c = y_v$$

And we could repeat this for any number of dimensions. So, spatial utility functions have exactly the property we wanted: we can cleanly optimize them, and the unique maximum of a voter's utility is exactly at that voter's location, which we call their ideal point.

Exercise: The result here contains an extremely ugly contradiction.[10] If $x_c = x_v$ and $y_c = y_v$, then the denominator of the fraction is 0. This is actually a deep hole in the simple spatial optimisation idea. One way that we could work around this is by instead constructing the problem so that, as the candidate's position *approaches* the voter's ideal point, the voter increasingly likes the candidate, so that the denominator is infinitesimal but nonzero. Since we have just been studying limits and infinite series which are just like that idea, try to write out a version of spatial utilities that accomodates this idea. **Answer:** A sufficient answer is any idea which uses a limit to define a series which monotonically decreases in the distance between the points.

Note: Of course this isn't actually the utility of a vote – we also need to consider how much we like the other candidates, and how likely we are to change the result of the election.

(2) **Optimisation in games**:
Last time we saw probably the simplest and most ubiquitous types of utilities, spatial utilities, which are a very common quantity that gets optimised. As part of thoroughly introducing the idea of optimising utilities, let's now look at probably the oldest example of

[10] Thanks to David Chung for first pointing this out to me.

optimising utilities, and one that still gets brought up as a workhorse/reference in political science: the Cournot Duopoly model.

Example: This is a classical Cournot Duopoly model.[11] I have written this example to be extremely similar to the example using $\max_{x,y} f(x,y) = -\frac{1}{2}x^2 - xy - y^2 - 3y$ in Iain's unconstrained optimisation lecture from Lecture 18.

↪Say that q_1 and q_2 are the quantities of some product produced by firms 1 and 2.

↪Set $P(Q) = a - Q$ is the "market-clearing price" (equilibrium price, the price at which quantity supplied exactly equals quantity demanded), for the aggregate quantity on the market $Q \equiv q_1 + q_2$

↪Say also that the cost to firm $i, i \in \{1,2\}$ of producing quantity q_i is $C_i(q_i) = cq_i$ where c is just some constant.

↪Finally, say that the players move simultaneously.

Let's start by writing this game out in the standard format of strategic games: by identifying the players, the strategies, and the payoffs. When a game is written out in this manner it is called a "normal form game". We have:

↪Players: firms 1 and 2

↪Strategies: either firm can choose from the continuous strategy set $S_i = [0;\infty]$.

↪Payoffs: say that each firm's payoff is simply its profit. Then the payoff $u_i(s_i, s_j)$ is the price of the goods minus the cost of producing them, which is given as follows:

$$u_i(q_i, q_j) = q_i P(q_i + q_j) - cq_i$$

$$u_i(q_i, q_j) = q_i(a - (q_i + q_j)) - cq_i$$

$$u_i(q_i, q_j) = q_i a - q_i^2 - q_i q_j - cq_i$$

Now, to find the equilibrium, we want to figure out how each player i can maximize their utility u_i. So let's start by checking the first order condition for either player i:

$$0 = \frac{\partial}{\partial q_i}(q_i a - q_i^2 - q_i q_j^* - q_i c)$$

[11] The meat of this example is entirely taken from Gibbons (1992: p. 15-17), but I have dramatically changed the formatting and wording, and added some commentary. A related exposition to this one is in my 681 section notes from Winter 2019, available on my website.

$$0 = a - 2q_i - q_j^* - c$$

$$2q_i = a - q_j^* - c$$

$$q_i^* = \frac{1}{2}(a - q_j^* - c)$$

So, we obtain the following optimal quantity choices for the two firms:

$$q_1^* = \frac{1}{2}(a - q_2^* - c)$$

and

$$q_2^* = \frac{1}{2}(a - q_1^* - c)$$

Now we can just solve the pair of equations by plugging one into the other:

$$q_1^* = \frac{1}{2}(a - \frac{1}{2}(a - q_1^* - c) - c)$$

$$2q_1^* = a - \frac{1}{2}a + \frac{1}{2}q_1^* + \frac{1}{2}c - c$$

$$2q_1^* - \frac{1}{2}q_1^* = \frac{1}{2}a - \frac{1}{2}c$$

$$\frac{3}{2}q_1^* = \frac{1}{2}a - \frac{1}{2}c$$

$$q_1^* = \frac{a - c}{3}$$

And, by the symmetry of the equations, we find similarly

$$q_2^* = \frac{a - c}{3}$$

Was the first order condition indeed sufficient to find the maximum? Let's check the Hessian matrix, which is:

$$D^2 u_i = \begin{bmatrix} \frac{\partial^2 u}{\partial q_1^2} & \frac{\partial^2 u}{\partial q_1 \partial q_2} \\ \frac{\partial^2 u}{\partial q_2 \partial q_1} & \frac{\partial^2 u}{\partial q_2^2} \end{bmatrix}$$

$$D^2 u_i = \begin{bmatrix} -2 & -1 \\ -1 & 0 \end{bmatrix}$$

The determinant of the first leading principal minor is negative since

$$|A_1| = -2$$

And determinant of the second leading principal minor is negative since

$$|A_2| = (-2 \cdot 0) - (-1 \cdot -1)$$

$$|A_2| = -1$$

So the Hessian is indefinite, which means that this isn't in general sufficient to find the global maximum (positive definite would mean a minimum, negative definite would mean a maximum, indefinite means a saddle point). It turns out that the function is concave down if $q_j < a - c$, which is true so long as the players are playing the equilibrium strategy. How can we interpret this? In retrospect, it isn't too surprising, because we only maximized the function *assuming that the other player is also maximizing their function*, which is the core conceit of strategic interactions. If the other player for whatever reason does not maximize their utility function, then we have a completely different situation where we might have the opportunity to obtain a higher utility value.

Since this is the maximum of both players' payoffs, we say that this is the unique Pure Strategy Nash Equilibrium of the classical Cournot Duopoly game. But what's more important is that again we have seen that payoffs in a strategic interaction are just a function that can be maximized like any other function.

10 Section 10: 2020 December 1

Topics:

- 598 in mathematical context[1]
- Math topics outside of 598[2]

After this class, I expect every student to be able to:

- know where to look as a first step upon encountering any of these math topics out in the wild!

BACK TO TABLE OF CONTENTS

This section has two purposes:

(i) To acknowledge that there are mathematical topics that you might encounter in the political science literature that we did not have time to cover in 598, and

(ii) To give you a starting point for how to learn about those topics when you see them.

On the topic of (i), however, I want to be crystal clear that what I am talking about today are all *niche* topics. 598 covers by far the bulk of the math that is used in political science, and what we'll see throughout this section is that there are strong connections between topics learned in 598 and every other domain of applied math.

(1) **598 in mathematical context**:

We could roughly understand the material in 598 as follows. A very common dichotomy in discussing how math is studied is to talk about the difference between algebra and analysis. Along these lines, 598 contains more or less three threads.

$\hookrightarrow$ Algebra topics: Vectors $\to$ matrices $\to$ eigenvalues and eigenvectors $\to$ vector spaces

$\hookrightarrow$ Analysis topics: Functions $\to$ continuity $\to$ optimisation $\to$ fixed points

Some major topics not listed here, like utility functions, are examples of the topics above: for example, utility functions are functions (sometimes they are correspondences, which we also talked about a lot). We also studied a few niche topics in the foundations of math

$\hookrightarrow$ Foundations topics: Sets $\to$ sequences $\to$ limits

Of course, these boundaries are fuzzy – we very seriously studied compactness, which is a property of sets of numbers but is entirely about the existence of limits, and many other topics cross the boundaries like this. If you would like a more detailed idea of how everything fits together then I highly recommend the very great book Kline (1972).

(2) **Math topics outside of 598**:

For each topic that I mention in this section, I describe three things: what is the topic, how is it used in political science, and how does it relate to topics that you've learned in 598? I'm restricting myself to papers or books which are applications, **not** methodological innovations, and which were published within the past 10 years.[12]

[12] I impose an additional restriction. It's a sad reality that if you judged a methods class by the demographics in its bibliography and theorem names alone, you might reasonably conclude that the class is being taught in Vienna in the year 1890, both because of who was enabled to develop these ideas in the first place, and because of contemporary inequities over the past few decades in who has written a lot of the most-used math texts. So I take this opportunity to only provide references to papers where the first author is a woman, with emphasis on including work by women of colour.

Geometry (especially trigonometry):

Geometry is the study of shapes, trigonometry is specifically the study of triangles.

Recall: The basic trigonometric relations measure the properties of a triangle.

Example: Vector geometry, including some trigonometry ideas, is heavily used in text analysis. The simplest model is called "bag of words". First you pick a dictionary of words, like a standard English dictionary. Say that there are D words in the dictionary, then we construct a vector of length D. Then we count the number of times that each word in our dictionary occurs in the text document that we care about, and we record that number in the corresponding position of the vector. So if the word "platypus" occurs 700 times in our document, then the element of the vector that corresponds to the word "platypus" will contain the number 700. This is the simplest type of word vectorisation – it was the state of the art about 50 years ago.

Of course, the huge piece that's missing from this method is that it is extremely difficult to say anything about relationships between words, since we haven't captured how often words occur *in relation to one another*. So now let's think about ways of capturing pairs of words. Say that I have the two corpuses {I love cats} and {cats love my food}. Now consider the matrix that records how often each word is adjacent to each other word in our corpus:

	I	love	cats	my	food
I	0	1	0	0	0
love	1	0	2	1	0
cats	0	2	0	0	0
my	0	1	0	0	1
food	0	0	0	1	0

We can already make some neat conclusions just from looking at this table. Love is a word which is closely related to cats and closely related to me. Food is related to me, but not to cats. But here's the most interesting piece: imagine that you are an alien who doesn't speak any human language and I asked you to just look at the numbers associated with the word "I", namely the vector $\begin{bmatrix}0 & 1 & 0 & 0 & 0\end{bmatrix}$, and I asked you: what is the closest word to "I" in this mix, judging only by the vectors associated with them?

I think there are two answers: you could name the vector associated with "cats", which is $\begin{bmatrix}0 & 2 & 0 & 0 & 0\end{bmatrix}$, or you could name the vector associated with "my", which is $\begin{bmatrix}0 & 1 & 0 & 0 & 1\end{bmatrix}$. So your alien brain would have two hypotheses: either "I" and "my" mean something very similar, or else I am a cat. The idea is that in a sufficiently large corpus these ambiguities go away.

Now recall that a $k-$dimensional vector specifies a direction in $k-$dimensional space. Surprisingly, this means that a geometric measure of similarity between two lines can be used to measure the similarity between two words! If we imagine that all of the vectors meet at one point, then one such measure is the angle between each vector. As it happens, there are particularly efficient ways of finding the cosine of an angle. So if you ever venture into the world of text analysis, the baseline notion of similarity between two words is the "cosine similarity", which is simply the cosine of the angle between the vectors associated with any two words.

You might say that adjacency is a very simplistic way of defining similarity. Of course you're right – the biggest contemporary word vectorisation packages use extremely fancy machine learning methods to construct the vectors associated with words, based on very complicated ways of classifying their relative positions in a document. And the results they produce can be extremely sophisticated. The famous result used to advertise one of the major word embeddings (word vectorisation) packages a few years ago is that when you take the vector that this package associates with the word "king" and subtract the vector that it associates with the word "man", the closest vector out of all the words in its corpus to the result is the vector that it associated with the word "queen".

If you are interested in a political science example, the paper Rodman (2019), called *A Timely Intervention: Tracking the Changing Meanings of Political Concepts with Word Vectors*, asks how the use of words in political science drifts over time, and tests different text analysis packages against words that were known to have shifting meaning. This should be a very hard test for packages that take the meaning of the words to be fixed (a word usage from 1922 in my corpus should affect the procedure I described as much as a word usage from 2019), and it's a really good example of how powerful word embeddings can be.

Example: Plenty of formal models are extremely geometric. One example is the theory of Veto Players, invented by George Tsebelis. The model requires thinking about the size of various winsets in some policy space, which is the region of the space where a proposal would be able to obtain a sufficient number of votes to pass. So George's talks and papers often involve literally drawing geometric figures to find the winset, and a hypothesis can absolutely hinge on, say, the area of a hexagon. As a substantive example of how people apply this purely geometric idea to political science questions, I recommend Angelova et al. (2019), *Veto player theory and reform making in Western Europe.*

Of course, spatial voting logic is maybe the central theoretical idea in the history of quantitative political science, and the geometric problems here abound.

Connections and where to start: We did almost none of this in 598, although we did see the unit circle in some constrained optimisation examples. However, I believe that this is a topic that everyone encountered in detail in high school. All of the geometry that is every used in political science is much simpler than the topics we've been studying in 598, so as a starting point anything you learned from in high school classes is probably sufficient to remind you of the main properties of simple polygons.

Logic:

Logic studies the relationships between propositions, and whether or not those propositions are true.

Example: (Causal inference) Pure logic is a surprisingly fertile tool for applications and innovations in causal inference. We can see this just by thinking about the idea of a confounding variable, which is maybe the most basic problem statement in causal inference. If we have a situation where X causes Y and causes Z, while Y does not cause Z, then we might incorrectly conclude that Y causes Z simply because an unobserved value in X causes simultaneous changes in Y and Z. But if you replace "causes" with "implies", you will see that this is just a logic problem!

As a political science application, Falleti and Lynch (2009) give a really clean example in *Context and Causal Mechanisms in Political Analysis*: they survey and refine the idea of causal mechanisms in political science, which they say take statements of the form "If I, then O" ($I \to O$) and turn them into statements of the form "If I, through M, then O" ($I \to M \to O$). This sort of pure logic is still extremely productive in statistical thinking for political inference.

Connections and where to start: Believe it or not, logic and math are historically separate disciplines, which around the turn of the last century were deliberately unified against very active protestations by many logicians as well as many mathematicians. The main logical study that we made this semester was some special topics in set theory. As a starting place, I would recommend reading some texts specifically about logic in causal inference, rather than trying to crack open a pure logic textbook (although some people think the other way, and believe that the first step in learning how to do social science research designs should be reading a full introduction to logic). This is one of the most prominent conversations in statistics and social science, so it is easy to find good modern texts that engage with what exactly causality means in different domains.

Other calculus topics:

Two central calculus topics that we didn't cover are integration and ODEs.

We didn't cover integrals in 598, but you saw them plenty in math camp and 599.

Ordinary differential equations are derivatives of functions of a continuous parameter, which in scientific applications typically represents time, so it is the time-derivative of some function in the form $\frac{\partial}{\partial t} f(x,t)$. The idea is that these equations model the change in some function of the variable x as time progresses.

Example: ODEs are used whenever we care about how a value changes in time. A lot of these applications got filtered through Evolutionary Game Theory, which combines game theory together with ODEs to study the changes in animal populations over time. People try to bring these tools into political science every so often, and one example is in vote choice modeling: we can think about the number of supporters of one party versus another as a proportion of a population that evolves with time (I wanted to do this in my dissertation for a few months, but then I stopped believing in this particular application). A really fun paper along these lines is Guevara et al. (2018), *Evolution of Electoral Preferences for a Regime of Three Political Parties*, where the x in $\frac{\partial}{\partial t}f(x,t)$ represents the level of support for a given candidate, so $\frac{\partial}{\partial t}f(x,t)$ is the change in that support level over time. This study is called dynamical systems.

Connections: The standard text for learning ODEs is Tenenbaum and Pollard (1963), while the flavour of how they are used to study scientific phenomena is given really nicely by Strogatz (1994).[13]

Scientific computing:

One concern for empiricists or modelers is numerical linear algebra, the study of operations on matrices (which, remember, is how data is stored) which do not propagate errors. Such algorithms are called "stable". It turns out that many of these algorithms are based on creative matrix factorisations. By factorising a matrix we can stably apply operations to it that, in its un-factorised form, would introduce explosive errors. Relatedly, if you have a whole lot of data, then the speed of your algorithms can really matter too.

Example: You know an algorithm for row-reducing a matrix: Gauss-Jordan elimination. The problem is that if you instructed the computer to do exactly what you do when you row-reduce a matrix using Gauss-Jordan elimination, this algorithm is not stable. So there are stable modifications of Gauss-Jordan elimination which involve modified Gaussian elimination that relies on the LU factorisation.

Example: This work is often buried because algorithmic stability and efficiency aren't things you get to brag about in political science journals, sadly. I have spent a ton of time thinking about error propagation in algorithms in my own work, so I can definitely say that it matters if you are dealing with very small numbers. But an example of something that does get forefronted is eigendecompositions, which are very frequently used as an efficient and stable way of finding the eigenvalues of a matrix. A really nice example of a paper that heavily uses this idea to study opinion dynamics and for purposes like understanding the spread of fake news is Koprulu et al. (2016), *Battle of Opinions Over Evolving Social Networks*, using eigendecompositions.

Connections: We talked about two matrix decompositions in this class. One was the LU decomposition, $A = LU$, where L is lower triangular and U is upper triangular. The other was the eigendemposition, $A = X\Lambda X^{-1}$, where the columns in X are the eigenvectors of A and the diagonal extries of Λ are the corresponding eigenvalues. Probably my favourite textbook of all time is the standard text in numerical linear algebra: Trefethen and Bau III (1997), which to me is quite simply the gold standard of mathematical communication.

[13] For a more rigorous treatment I really like the book Hirsch et al. (2004). The combination of ODEs and game theory was achieved by John Maynard Smith in the topic called Evolutionary Game Theory. He has many books and articles on this really fun topic, which does find some extremely rare applications in political science.

Numerical linear algebra also has a pretty good Wikipedia page – at least, I think it's good, since I'm the one who wrote it! :)

Graph theory:

Roughly, the study of nodes connected by edges.

Network analysis is just an application of graph theory. In network analysis, we think of individuals as the nodes (say, people or countries) connected by edges that represent some relationship between those individuals (say, the opportunity to pass information to each other like on social media, or a trade agreement between countries). So we might say that countries A, B, and C form a network of countries we care about. Countries A and B each have a trade agreement with country C, but do not have a trading relationship with each other. So if we wish to model the passage of goods between the countries, then our network tells us that any goods which travel between countries A and B must travel through country C.

One great scholar who studies this is named Iain Osgood, and you can see it applied in his and Jieun Lee's paper, Lee and Osgood (2018), called *Exports, jobs, growth! Congressional hearings on US trade agreements.*

Connections: Graph theory is closely related to topology, which is the study of particular types of geometric objects. Since topologists often study the relationships of large number of points in a space, taking a set and applying structure to it to obtain some topological surface, this is the subfield that really gives us a rich understanding of topics like continuity and compactness. So while networks look very different from these ideas, you have actually studied much more related topics than you will think at first: just keep in mind that these topology topics are all about the relationships between sets of points. I don't have any specific introductory materials to recommend, but one of the primary inventors of network analysis, Mark Newman, regularly offers a graduate level class at the University of Michigan that students in our department have taken. If you have any interest in network analysis, you couldn't possibly ask for a better source to learn that topic from.

Complex analysis:

Complex analysis is the study of complex functions, which are functions of complex numbers: numbers of the form $a + bi$, where $a, b \in \mathbb{R}$ and $i \equiv \sqrt{-1}$.

Examples: Complex analysis is probably the math topic that has the highest ratio of prevalence in natural sciences to prevalence in social sciences, and I'm honestly not sure why that is. In natural sciences complex analysis is an exceptionally fundamental topic, whereas in political science we almost never encounter any topics that touch it even at the frontier of research. But there is one tool in complex analysis which I predict will inevitably gain in popularity at some point and find applications in political science because it is just such an extremely fundamental tool across all of science: Fourier Transforms. It turns out that any function which is integrable on some interval can be approximated by a sum of trigonometric functions. This is useful primarily because it can be much easier to apply various tools to a sum of, say, sine functions than it is to apply them to some arbitrarily messy functional form.

It has so far not seen extensive use in political science, but the main area where Fourier Analysis is used is in the study of time series data. One example is Chawsheen and Broom (2017), *Seasonal time-series modeling and forecasting of monthly mean temperature for decision making in the Kurdistan Region of Iraq*.

Connections: Maybe the nicest bridge to things we've mentioned previously is Euler's formula, which states that $e^{xi} = \cos(x) + i\sin(x)$ for some $x \in \mathbb{R}$. This is probably the single most famous equation in math (I guess the most famous equation is probably $1 + 1 = 2$) in the special case that $x = \pi$, which reduces to $e^{\pi i} = -1$, or the more ham-fisted but more popular form that you may have seen: $e^{\pi i} + 1 = 0$. We are close friends with e from our study of calculus, so the imaginary number system is only a short hop away.

References

Angelova, M., Bäck, H., Müller, W. C., and Strobl, D. (2019). Veto player theory and reform making in western europe. *European Journal of Political Research*, 57.

Chawsheen, T. A. and Broom, M. (2017). Seasonal time-series modeling and forecasting of monthly mean temperature for decision making in the kurdistan region of iraq. *Journal of Statistical Theory and Practice*, 11.

Downs, A. (1957). *An economic theory of democracy.* Harper.

Edwards Jr., C. H. (1973). *Advanced calculus of several variables.* Dover Publications.

Falleti, T. G. and Lynch, J. F. (2009). Context and causal mechanisms in political analysis. *Comparative Political Studies*, 42.

Gibbons, R. (1992). *Game theory for applied economists.* Princeton University Press.

Guevara, M. G. M., Rodríguez, H. V., Padilla, P. B. E., and Solís, J. L. G. (2018). Evolution of electoral preferences for a regime of three political parties. *Discrete Dynamics in Nature and Society*, 2018.

Hirsch, M. W., Smale, S., and Devaney, R. L. (2004). *Differential equations, dynamical systems & an introduction to chaos.* Elsevier.

Kline, M. (1972). *Mathematical thought from ancient to modern times.* Oxford University Press.

Koprulu, I., Kim, Y., and Shroff, N. B. (2016). Evolution of electoral preferences for a regime of three political parties. *IEEE.*

Lee, J. and Osgood, I. (2018). Exports, jobs, growth! congressional hearings on us trade agreements. *Economics & Politics*, 31.

Mathews, J. and Walker, R. L. (1970). *Mathematical methods of physics.* The Benjamin/Cummings Publishing Company.

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective.* MIT Press.

Rodman, E. (2019). A timely intervention: Tracking the changing meanings of political concepts with word vectors. *Political Analysis.*

Rosenthal, D., Rosenthal, D., and Rosenthal, P. (2014). *A readable introduction to real mathematics.* Springer.

Strogatz, S. (1994). *Nonlinear dynamics and chaos.* Westview.

Tenenbaum, M. and Pollard, H. (1963). *Ordinary differential equations: An elementary textbook for students of mathematics, engineering, and the sciences.* Dover.

Trefethen, L. N. and Bau III, D. (1997). *Numerical linear algebra.* Society for Industrial and Applied Math.