

1 January 8

1.1 Logistics

Within the three body problem is the entirety of this course.

“In mathematics you don’t understand things. You just get used to them.” - John von Neumann. (Bro von Neumann is so washed for this.) For context, von Neumann was challenged that he didn’t actually understand the method of characteristics. He was talking about this with his Dad, who just retired as a math prof from Ohio State, and he didn’t like it. Joel’s alternate form: “In mathematics it takes time for ideas and examples to sink in.”

Click here for course website link.

Content: Parts of Chapters 10-12, 14-16 of textbook by Ascher and Greif (online, free). Topics: Interpolation, approximation (Ch 10 - 12); Differentiation, Integration, ODE’s [PDE’s] (Ch 14-16). Back in the day, 303 was meant as a followup to 302, but not anymore, so may be some repeated material (norms and condition numbers, but not the point of this course anyway).

Discussion: please post to piazza page. If this fails, please email to jf@cs.ubc.ca with subject CPSC 303.

Grading: $(10\%) \max(h, m, f) + (35\%) \max(m, f) + (55\%) f$ where h is homework, m is midterm, f is final. So technically, can ace final and ace course. Because back in the day, Joel knew someone who couldn’t attend any classes, but got 100% in the final only to get C+ in the course. There is a phenomena where if someone gets 100 on the midterm, stops doing homework. But really, these assessments are good preparation and indicators of where you stand in the course.

Please sign up for piazza and gradescope through canvas.ubc.ca (especially gradescope).

Homework: Set Thursday 11:59pm and due Thursday 11:59pm. There is both individual and group homework. Group homework: at most 4 people, and will cover most material; only submit one. Individual homework: These are the types of things he wants to make sure everyone can do, and will be like the things on exams; you must write up your own solution even if you work with others.

1.2 Intro to ODE’s

1.2 and 4.2 (norms), 14.2 (differentiation), 16.1 and 16.2 (ODE’s).

He typically begins courses with the most difficult stuff the course will get. This is tough, but not the worst. Now, this course only requires two terms of calculus, but with how pertinent ML is now, most people are taking multivariable calc anyway. We will get some idea of what to expect and some intuition, but will revisit later in the course. The reason the emphasis is on ODEs and not PDEs is because the general theory for ODEs really applies to all of them, even if you have to solve differently. Families of PDEs have their own properties that have to be studied separately.

We are heading towards Ordinary Differential Equations (Ch. 16).

1.2.1 Absolute vs. Relative Error

If $v \in \mathbb{R}$ is an approximation to $u \in \mathbb{R}$, then absolute error (in v) (as an approximation to u) is $|u - v|$, and the relative error is $\frac{|u - v|}{|u|}$. The same works in \mathbb{R}^n (or any normed vector space):

$$\|\vec{u}\|_2 = \|(u_1, \dots, u_n)\|_2 = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2}$$

We also use $\|\vec{u}\|_1 = |u_1| + \dots + |u_n|$ and $\|\vec{u}\|_{\max} = \|\vec{u}\|_{\infty} = \max_{1 \leq i \leq n} |u_i|$. The absolute error in \vec{v} as an approximation to \vec{u} is $\|\vec{u} - \vec{v}\|_p$ and the relative error is $\frac{\|\vec{u} - \vec{v}\|_p}{\|\vec{u}\|_p}$ where $p = 1, 2, \infty$.

1.2.2 Taylor’s Theorem (p. 5)

Theorem 1. For $f: (a, b) \rightarrow \mathbb{R}$ where f is $k + 1$ differentiable (so $f^{(k+1)}(x)$ exists in (a, b)), for some $x_0, x_0 + h$ that lie in (a, b) ,

$$f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2} f''(x_0) + \dots + \frac{h^k}{k!} f^{(k)}(x_0) + \text{error}$$

where the error is $\frac{h^{k+1}}{(k+1)!} f^{(k+1)}(\xi)$ where ξ is between x_0 and $x_0 + h$.

Remark 1. h can be negative.

We'll learn how to approximate derivatives, and then ODE solution.

2 January 10

Housekeeping:

- HW1 will be assigned on Jan 11, due on Gradescope on Jan 18
- Access Gradescope via Canvas
- Today: Separable ODE's (see, e.g. CLP 2, Appendix D of Ascher and Grief)

Last time: Ch 1: "Reviewing" terminology. Ch 4: $\|\vec{u}\|_2 = \sqrt{u_1^2 + \dots + u_n^2}$ for $\vec{u} \in \mathbb{R}^n$. We saw classes of functions, Taylor Series, and ODE's. An ODE is where the derivatives are a function of the same variable?? So $y' = f(t, y)$. PDEs on the other hand have partial derivatives $h = h(t, x_1, \dots, x_n)$, and maybe something like the heat equation $\frac{\partial h}{\partial t} = -\Delta_{x_1, \dots, x_n} h$. There might be a course on the foundations of elliptic PDEs, parabolic PDEs, etc., but won't be the focus here.

And then he talks about a sketch of proof of Taylor's, except I literally did it this morning. Essentially, we use MVT to get better approximations (not rigorously). We have

$$f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2}f''(x_0) + \frac{h^3}{6}f'''(\xi_1)$$

$$hf'(x_0) = f(x_0 + h) - f(x_0) - \frac{h^2}{2}f''(\xi_2)$$

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h} - \frac{h}{2}f''(\xi_2) = \frac{f(x_0 + h) - f(x_0)}{h} + \text{Order}(h)$$

where $O(h)$ is some function (depends on f, x_0). We have $|\text{Order}(h)| \leq \frac{h}{2}M_2$ where M_2 is bound on f'' in the interval $[x_0, x_0 + h]$. By definition, $f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$. All this is really taken from [A&G], Ch. 14, Sections 1,2.

We will continue our discussion of derivatives, but first, some useful notation. If $a < b \in \mathbb{R}$, $C[a, b] := \{f: [a, b] \rightarrow \mathbb{R} \text{ such that } f \text{ is continuous}\}$. When $k \in \mathbb{N}$,

$$C^k(a, b) = \{f: (a, b) \rightarrow \mathbb{R} \text{ such that } f \text{ has } k \text{ continuous derivatives } \forall x \in (a, b)\}$$

, and similarly for $C^k[a, b]$. $C^\infty(a, b)$ is the set of f that has derivatives of all order.

Definition 1 (Real Analytic Functions).

$$C^\omega(a, b) := \{f: (a, b) \rightarrow \mathbb{R} \text{ such that for all } x_0 \in (a, b), f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2!}f''(x_0) + \dots\}$$

A function $f: (a, b) \rightarrow \mathbb{R}$ is real analytic when $f \in C^\omega(a, b)$.

Real analytic is a stronger condition than C^∞ , so $C^\omega \subset C^\infty \subset \dots \subset C^2 \subset C^1 \subset C^0$. [Hmm... I would like a better definition of convergence here ff]

For example, e^x ff he scrolled.

2.1 Start ODE

Simple ODE [A&G]:

$$y' = f(t, y)$$

where we use the notation $y' = \frac{dy}{dt} = \dot{y}$. (Caution: math textbooks typically use $y' = \frac{dy}{dx} = f(x, y)$.)

To solve for $y = y(t)$, we are given an "initial condition", that is we have $y_0, t_0 \in \mathbb{R}$ and impose $y(t_0) = y_0$.

We expect a unique solution. Say we are given $A \in \mathbb{R}$, and find a y that satisfies

$$y'(t) = Ay(t)$$

We claim that $y(t) = e^{At}C$ is a solution. We can verify: $y'(t) = (e^{At}C)' = (Ae^{At})C = Ay(t)$. So we have solved it by shamelessly guessing.

We can plug in our t_0 and y_0 which fixes $C = y_0 e^{-At_0}$. Then

$$y(t) = y_0 e^{A(t-t_0)}$$

So when A is big enough (greater than 0), we have exponential growth.

Is this solution unique? Can we simply guess and it provides the only solution? More on Friday.

3 January 12

Today:

- ODE: $y' = y$ “isoclines”
- ODE's of the form $y' = f(y)$
- Later $\vec{y}' = \vec{f}(t, \vec{y})$, system of m ODE's where $\vec{y} = \vec{y}(t): \mathbb{R} \rightarrow \mathbb{R}^m$ and $\vec{f}(t, \vec{y}): \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$.
- Examples; $y' = y^2$, $y' = |y|^{1/2}$.

Question, what could possibly go wrong?

Solve $y' = f(y)$, $y(t_0) = y_0$. Then

- If f continuous near $y = y_0$ then a solution exists locally
- Moreover, if f is Lipschitz (or differentiable) near $y = y_0$, the local solution is unique
- If f is analytic near $y = y_0$, the unique local solution is analytic
- If $|f(y)| \leq Ky$ for some K constant, then there is a global solution

We will start looking at Matlab next week: we will specifically be looking for how it breaks. He doesn't cover as much content as other people do in 303, but more in depth. “I just became a grandfather last week, so I get to say back in my day we always had to make our own software.” So we always knew how it broke. Now we have to figure out why other people's code breaks.

Recall last time, we were looking at $y' = Ay$, $y(t_0) = y_0$ where $y: \mathbb{R} \rightarrow \mathbb{R}$. Isocline picture: slopes at unit points in the $y-t$ plane. And then we guessed an answer last time: $y(t) = e^{A(t-t_0)}y_0$. Now, if we had a theorem that said this was unique, we'd be done.

We can solve this with integration: $y' = y \implies \frac{dy}{y} = dt \implies \int \frac{1}{y} dy = \int dt$, and so $\ln(y) + c_1 = t + c_2$ hence $y = e^{t+c}$. Now with t_0, y_0 , we can determine our constant to get $y(t) = e^{t-t_0}y_0$. However, this method is not foolproof. Our solution could blow up to infinity. Consider $y' = y^2$ and $y(1) = 1$. And then $\frac{1}{y^2} dy = dt \implies \frac{-1}{y} = t + c$. Then, $y = \frac{-1}{t+c}$, and plugging in the initial conditions, we get $c = -2$. So $y = \frac{1}{2-t}$. Singularities can happen, as $y(t) \rightarrow \infty$ as $t \rightarrow 2$.

What about $y' = |y|^{1/2}$? What could possibly go wrong? When $y > 0$ we have $y' = y^{1/2}$. Solving in the way we did before, we can get $y^{1/2} = \frac{1}{2}(t+C)$ so $y(t) = \frac{1}{4}(t+C)^2$. Note $y(-C) = 0$... but our slope should always be increasing, yet is 0 at a point? ff idk Now when $y < 0$, we can find (with the method as before) that $y = -\frac{1}{4}(t+C)^2$. So we have piecewise function of y depending on if $y > 0$ or $y < 0$.

Is this a unique solution? Actually, let $a < b$. Then

$$y(t) = \begin{cases} \frac{1}{4}(t-b)^2 & b \leq t \\ 0 & a \leq t \leq b \\ -\frac{1}{4}(t-a)^2 & t \leq a \end{cases}$$

is also a solution. The bad situation occurs when $y = 0$. We could stay “arbitrarily” long at $y = 0$, even if to the right and left are parabola! We will continue looking at $y' = |y|^{1/2}$ next time.

4 January 17

Today's outline:

- Euler's method
- MATLAB and Euler's method
- Plugging in $y' = Ay, y(t_0) = y_0$

4.1 Euler's Method

Say we are given $y' = f(t, y)$ (or $y' = f(y)$, or $\vec{y}' = \vec{f}(t, \vec{y})$, etc.) where f is a function, and we have) initial value $y(t_0) = y_0, y_0, t_0 \in \mathbb{R}$. We have the approximation

$$y'(t) \approx \frac{y(t+h) - y(t)}{h}$$

for small h , since $y'(t) = \lim_{h \rightarrow 0} \frac{y(t+h) - y(t)}{h}$. But

$$y'(t) \approx \frac{y(t+h) - y(t-h)}{2h}$$

is a much better approximation (usually). Rearranging gives $y(t+h) \approx y(t) + hy'(t)$; in fact, by Taylor's theorem, $y(t+h) = y(t) + hy'(t) + \frac{h^2}{2}y''(\xi)$ for some ξ between t and $t+h$. Substituting f , we have

$$y(t+h) = y(t) + hf(t, y) + \frac{h^2}{2}y''(\xi)$$

So for small h ,

$$y(t+h) \approx y(t) + hf(t, y)$$

(we will ignore the final term for now, but will be useful later for calculating error). This last approximation is the essence of Euler's method.

So given $t_0 =$ initial time and $y_0 =$ initial value,

Step 1. Start with $y(t_0) = y_0$. [Pick a value of h , smaller the better (usually)]

Step 2. $y(t_0 + h) = y_0 + hf(t_0, y(t_0)) := y_1$

Step 3. $y(t_0 + 2h) = y_1 + hf(t_1, y_1) := y_2$ where $t_i = t_{i-1} + h = ih + t_0$.

Step n . Repeat

Actual ODE solvers (like in MATLAB) change h based off of f (especially makes h small when f is very large or f is changing quickly). When we saw the three body problem, we had error when two bodies got close because the gravitational force got so big it couldn't make h small enough to figure out what was going on.

If $y' = 2y$ and given y_0, t_0 , recall that the exact solution was $y(t) = e^{2(t-t_0)}$. With the numerical approximation, we get $y_1 = y_0 + h(2y_0) = y_0(1 + 2h)$ and $y_2 = y_1 + h(2y_1) = (1 + 2h)y_1$. So $y(t_i) = y(t_0 + ih) = (1 + 2h)^i y_0$. Now say we fix a t_{end} , perhaps N steps and so $t_{\text{end}} = t_0 + Nh$. Then $h = \frac{t_{\text{end}} - t_0}{N}$. So

$$y(t_{\text{end}}) = (1 + 2h)^N y_0 = \left(1 + \frac{2(t_{\text{end}} - t_0)}{N}\right)^N y_0 = \left(1 + \frac{\text{something}}{N}\right)^N y_0$$

Hmm... this looks a lot like a definition of e . As $N \rightarrow \infty$, we have $e^{\text{something}} y_0 = e^{2(t_{\text{end}} - t_0)} y_0$. We will see this in MATLAB. On the homework, we will see MATLAB not working too well for $y' = |y|^{1/2}$.

And then he shows us his MATLAB for Euler's method. MATLAB is quirky. The first function you define in a file will always be run when you call the file, and it will assume the other functions in the file are not meant to be run globally.

5 January 24

Topics to finish:

- Look at MATLAB for a bit (Solutions to HW2 and HW1-ish)
- Vandermonde matrices and linear algebra without linear algebra
- Exponentiation and eigenpairs and norms
- Higher order versions of Euler's method
- More celestial mechanics

Some quirky MATLAB: Something good to know is putting a `;` at the end of a line to suppress the output (so we don't see giant vector). `1e-200` is defined as its own thing, but `1e-400` is considered `0`. `1/0` gives `Inf` and `-1/0` gives `-Inf`. But `-Inf + Inf` gives `NaN` (not a number). But all this follows the IEEE standard. MATLAB also calls the first function in a file as the name of a file when called elsewhere. We will look more at MATLAB later when we try to break it.

5.1 Vandermonde matrices

Given

$$A = \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix}$$

Does this have a unique solution?

The following are equivalent:

1. $A \begin{bmatrix} c_0 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} y_0 \\ \vdots \\ y_n \end{bmatrix}$ has a unique solution for all y_0, \dots, y_n .
2. A is invertible
3. $\det(A) \neq 0$
4. A is of rank $n + 1$

We have actually seen Vandermonde matrices in HW1, just disguised. We were looking at 3 point schemes. We took a bunch of these formulas:

$$\begin{cases} f(x_0 + L) = \cdots \\ f(x_0) = \cdots \\ f(x_0 - h) = \cdots \\ f(x_0 + 2h) = \cdots \\ f(x_0 + \sqrt{2}h) = \cdots \\ f(x_0 + rh) = \cdots \end{cases}$$

(where r is a real number, often an integer, often $0, 1, 2, -1, -2, \dots$), and used the formula

$$\begin{aligned} c_0 \cdot f(x_0 + r_0 h) &= c_0 \left(f(x_0) + r_0 h f'(x_0) + r_0^2 \frac{h^2}{2} f''(x_0) + r_0^3 \frac{h^3}{3!} f'''(x_0) + O(h^4) \right) \\ c_1 \cdot f(x_0 + r_1 h) &= c_1(\dots) \\ c_2 \cdot f(x_0 + r_1 h) &= c_2(\dots) \\ c_3 \cdot f(x_0 + r_1 h) &= c_3(\dots) \end{aligned}$$

and we try to rearrange and change c_i and derive a scheme to get $0f(x_0) + 1hf'(x_0) + 0\frac{h^2}{2}f''(x_0) + 0\frac{h^3}{3!}f'''(x_0)$. I.e., we are looking for c_0, c_1, c_2, c_3 such that

$$\begin{aligned} 0 &= c_0 + c_1 + c_2 + c_3 \\ 1 &= r_0c_0 + r_1c_1 + r_2c_2 + r_3c_3 \\ 0 &= r_0^2c_0 + r_1^2c_1 + r_2^2c_2 + r_3^2c_3 \\ 0 &= r_0^3c_0 + r_1^3c_1 + r_2^3c_2 + r_3^3c_3 \end{aligned}$$

In homework 1, we saw this was solving a vandermonde matrix (anything that is of the form of the matrix below):

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ r_0 & r_1 & r_2 & r_3 \\ r_0^2 & r_1^2 & r_2^2 & r_3^2 \\ r_0^3 & r_1^3 & r_2^3 & r_3^3 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

(Observation: this looks like interpolation.)

Theorem 2. *The above system has a unique solution.*

Lemma 1. *Say that $p: \mathbb{R} \rightarrow \mathbb{R}$, $p = p(x)$ is a polynomial and p has $n+1$ distinct roots, x_0, \dots, x_n . Then*

1. $p(x) = (x - x_0)(x - x_1) \dots (x - x_n)q(x)$ where $q(x)$ is a polynomial. (Bruh people in this class are actually so stupid, none of them recognize factoring).
2. $p'(x)$ has n distinct roots between x_0 and x_n (assuming $x_0 < x_1 < \dots < x_n$).
- (a). $p''(x)$ has $n-1$ distinct roots between x_0 and x_n
- (b). $p'''(x)$ has $n-2$ distinct roots between x_0 and x_n

To prove all of these, just done by Rolle's theorem

Theorem 3 (Rolle's Theorem). *If $f \in C^0[a, b]$ (continuous on $[a, b]$), $f(a) = f(b) = 0$, and f is differentiable on (a, b) , then $f'(\xi) = 0$ for some $a < \xi < b$.*

(Bro why is no one putting up their hand to haveing SEEN Rolle's theorem before, I swear there are too many bums in this course.) "How many people have seen a proof of this? Yeah, you would do this in Math 320." (Loewen could never) But this is quite intuitive anyway.

We do get a useful fact from the lemma:

Corollary 1. *If $p = p(x)$ is a polynomial of degree $\leq n$, $p(x) = c_0 + c_1x + \dots + c_nx^n$ and p has $n+1$ distinct roots, then $p = 0$*

Remark 2. $p(x) = x^2 + 1$, then p has no real roots. But $p'(x) = 2x$ has a real root, so the lemma doesn't go the other way (i.e. $q(x) = 0$ from bullet point 1.).

Claim: the system

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = 0$$

has $c_1 = c_2 = \dots = c_n = 0$ as its unique solution. Why? This just says $p(x) = c_0 + c_1x + c_2x^2 + \dots + c_nx^n$ satisfies $p(x_0) = 0, p(x_1) = 0, \dots, p(x_n) = 0$, so p has $n+1$ distinct roots, hence by our corollary, $p = 0$. (Important in the definition of a Vandermonde matrix is that x_0, x_1, \dots, x_n are distinct, otherwise not invertible).

Homework: $e \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^t = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$ and $e \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^t = \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}$ (meant to get to this today).

6 January 26

Coming soon (later, we're putting a pin in it):

- More ODE's:
 - Central force: $m\vec{x}'' = -\frac{\vec{x}}{|\vec{x}|^3}g(|\vec{x}|)$,
 - Newton's law: $g(r) = \frac{1}{r^2}$
 - ff
- ff

Today:

- Higher Order ODE solvers. For now, Euler's method and Explicit trapezoidal
- More on e^{At} , norms, convergence, etc.s
- Where do non-diagonalizable matrices come from? One place: recurrences

If $y' = f(t, y)$, $y(t_0) = y_0$, we can rewrite this as

$$y(t) - y_0 = \int_{s=t_0}^{s=t} f(s, y) ds$$

This is the integral form of " $y' = f(t, y)$ ". This is actually a preferred form for many applications, since we don't know if y is differentiable so we don't assume it. Rename: $f(s) = f(s, y)$ (or are we assuming??) Then $y' = f(t)$, $y(t_0) = y_0$, then $y(t) - y(t_0) = \int_{s=t_0}^{s=t} f(s) ds$. We want to use a trapezoid to approximate the area under the curve, since it is a better approximation than a rectangle.

We have $y(t_0 + h) = y(t_0) + hy'(t_0) + O(h^2)$. Let $t_i = t_0 + ih$. We want y_1 to approximate $y(t_1)$. $y(t_1) = y(t_0 + h) \approx y(t_0) + hy'(t_0) + O(h^2)$. So $y_{i+1} = y_i + hf(t_0, y_0)$. So y_{i+1} approximates $y(t_{i+1})$ (uses y_i , approximates $y(t_i)$).

How do we improve? $y(t+h) = y(t) + h\frac{y'(t)+y'(t+h)}{2}$, which is a better approximation for y' . Trapezoidal: $\frac{y(t+h)-y(t)}{h}$ to $y'(t + \frac{h}{2})$. (I have no clue what he is doing.)

So first, we have $y(t_0), h$. Using Euler, we get $y(t_1) = y(t_0 + h) \approx y(t_0) + hy'(t_0) = y(t_0) + hf(t_0, y_0)$. $y_1 \leftarrow y(t_0) + hf(t_0, y_0)$.

Trapezoidal (Explicit) Scheme: Let $Y = y(t_0) + hf(t_0, y_0)$. And $y(t_0 + h) \approx y(t_0) + h\left(\frac{f(t_0, y_0) + f(t_1, Y)}{2}\right)$. The Y is like the otherside of the trapezoid.

6.1 Pre-interpolation

Before we get to interpolation, we need to get back at matrices, and what norms of matrices are.

6.1.1 Matrix exponential

Question: What do we mean by

$$e^A = I + A + \frac{A^2}{2} + \dots$$

If $\vec{x} \in \mathbb{R}^n$, then $\|\vec{x}\|_2$ or $\|\vec{x}\|_{L^2}$ is $\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$ (I feel like this should be ℓ^2).

Definition 2 (Matrix norm L^2). If $M \in \mathbb{R}^{n \times n}$, i.e. an $n \times n$ matrix, real entries, then we define

$$\|M\|_{L^2} = \max_{\vec{x} \neq \vec{0}} \frac{\|M\vec{x}\|_2}{\|\vec{x}\|_2}$$

So $\|M\|_{L^2}$ is the smallest real B such that $\|M\vec{x}\|_{L^2} \leq \|\vec{x}\|_{L^2} B$.

Example: $\vec{x} \in \mathbb{R}^n$, $n = 1$, $\vec{x} = x_1 \in \mathbb{R}^1$ so $A \in \mathbb{R}^1$. Say $A = [3]$. $\max\left(\frac{\|A\vec{x}\|_{L^2}}{\|\vec{x}\|_{L^2}}\right)$. In this case, $\vec{x} = (x_1)$, so $\|\vec{x}\|_{L^2} = \sqrt{x_1^2} = |x_1|$. So $A \in \mathbb{R}^{|x|}$??? what is this notation, $A = [a]$, so $\|A\|_{L^2} = |a|$. ff

We also have other norms: $\|\vec{x}\|_{\max} = \|\vec{x}\|_{\infty} = \max(|x_1|, |x_2|, \dots, |x_n|)$. $\|A\|_{L^{\infty}} = \max_{\vec{x} \neq 0} \frac{\|A\vec{x}\|_{\infty}}{\|\vec{x}\|_{\infty}}$.

Now look at 2×2 matrices. Say

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$\text{size}(A) = \max(|a_{11}|, |a_{12}|, |a_{21}|, |a_{22}|)$. Claim: $\|A\|_{\infty} = \|A\|_2 = \text{size}(A)$, and so it doesn't matter what norm we pick.

What is $I + A + \frac{A^2}{2} + \frac{A^3}{3!} + \dots$. What does convergence mean? The r th term is $I + A + \frac{A^2}{2} + \dots + \frac{A^r}{r!}$. So we want to measure convergence:

$$\lim_{s, r \rightarrow \infty, s > r} \left(\frac{A^{r+1}}{(r+1)!} + \dots + \frac{A^s}{s!} \right) \rightarrow 0$$

(I'm assuming this is Cauchy). With a given metric, say 2, this means

$$\lim_{s, r \rightarrow \infty, s > r} \left\| \frac{A^{r+1}}{(r+1)!} + \dots + \frac{A^s}{s!} \right\|_2 \rightarrow 0$$

(but it is the same with the other metrics). (I'm curious, how many of our properties from \mathbb{R}^n carry over to these matrix spaces, like we always have equivalence of these L^2, L^{∞} norms?). We have e^{At} where $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then computing our terms,

$$\begin{aligned} I &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ A &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ A^2 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ A^3 &= A^2 A = I A = A \end{aligned}$$

So for every powers, it is I , and if it is an odd power, it is A . Then we have

$$e^{At} = I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \dots = \begin{bmatrix} 1 + \frac{t^2}{2} + \frac{t^4}{4!} + \dots & t + \frac{t^3}{3!} + \frac{t^5}{5!} + \dots \\ t + \frac{t^3}{3!} + \frac{t^5}{5!} + \dots & 1 + \frac{t^2}{2} + \frac{t^4}{4!} + \dots \end{bmatrix} = \begin{bmatrix} \cosh(x) & \sinh(x) \\ \sinh(x) & \cosh(x) \end{bmatrix}$$

We'll do more diagonalization things next time.

7 January 29

Today:

- e^{At} where $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$
- Other reasons to write $A = S \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} S^{-1}$ other than e^{At} , namely recurrences and $A^n, n = 0, 1, \dots$
- Examples of
 - (a). Diagonalizable A
 - (b). Defective A (in recurrences) ("not enough eigenvectors")... (not diagonalizable, because it does not have a complete basis of eigenvectors, some duplicates)

What is a 2×2 defective matrix? Means a (real) matrix that is not diagonalizable over \mathbb{C} . Diagonalizable:

$$A = S \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} S^{-1}$$

hence

$$f(A) = S \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{bmatrix}$$

The rotation matrix will have real entries, but complex (unit length) eigenvalues. So it isn't defective.

While with projection onto 2-dimensional plane, you'll have eigenvalue 1 with multiplicity 2.

So we have gotten some intuition of eigenvalues. Greater than 1, it is stretching it along the corresponding eigenvector, and less than 1, it is shrinking it. So iterated powers either make it go to infinity, or to 0.

7.1 Recurrence Relations and Finite Precision

7.1.1 Fibonacci numbers:

The Fibonacci numbers are defined as $F_{n+2} = F_{n+1} + F_n$ with the initial conditions $F_1 = F_2 = 1$. So our sequence is $1, 1, 2, 3, 5, 8, 13, 21, 34, \dots$. We can also extend this backwards for $F_0, F_{-1}, F_{-2}, \dots$ with the recurrence relation $\dots, -8, 5, -3, 2, -1, 1, 0, 1, 1, \dots$; as you can see, backwards follows a similar pattern.

We can use some linear algebra here. Since, we have $F_{n+2} = F_{n+1} + F_n$ and $F_{n+1} = F_{n+1}$, we can write

$$\begin{bmatrix} F_{n+2} \\ F_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix}$$

We can iterate backwards until we get to our initial conditions, $\begin{bmatrix} F_2 \\ F_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} F_2 \\ F_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, so

$$\begin{aligned} \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} &= A \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} \\ &= A \cdot A \cdot \begin{bmatrix} F_{n-1} \\ F_{n-2} \end{bmatrix} \\ &\vdots \\ &= A^n \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \end{aligned}$$

See how this is similar to what we did with Euler's method, except the "dynamics" dictate that we take the power instead. There is a strong connection to Euler's method, $\vec{y}' = A\vec{y}$. We have

$$\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = A^n \begin{bmatrix} F_1 \\ F_0 \end{bmatrix}$$

How to diagonalize, at 2×2 matrix? Recipe: look for \vec{v} such that there is some λ where $A\vec{v} = \lambda\vec{v} = \lambda \cdot I\vec{v}$. So $(A - \lambda I)\vec{v} = 0$. So we look for λ s such that $A - \lambda I$:

- (a). has non-zero vector in its nullspace
- (b). $\det = 0$
- (c). $\text{rank} < n$ (for $n \times n$)

(d). not invertible

(e). etc. (your favourite condition)

We can solve (obviously not the move for higher dimensions):

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

So we are solving $(1-\lambda)(-\lambda) - (1)(1) = 0$, or $\lambda^2 - \lambda - 1 = 0$. The solutions to this are the golden ratio and its “conjugate”: $\lambda = \frac{1 \pm \sqrt{5}}{2}$. So given that $\lambda = \frac{1+\sqrt{5}}{2}$, \vec{v} such that

$$\begin{bmatrix} 1 - \frac{1+\sqrt{5}}{2} & 1 \\ 1 & -\frac{1+\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

So we want $(1 - \frac{1+\sqrt{5}}{2})v_1 + v_2 = 0$. If we take $v_2 = 1$, then $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} ? \\ 1 \end{bmatrix} = \frac{1+\sqrt{5}}{2} \begin{bmatrix} ? \\ 1 \end{bmatrix}$.

Let's look at a simpler recurrence. $F_{n+1} = 10F_n$. Then $F_n = 10^n F_0$. Just like an ODE, if we have an initial condition, reduces to e^{At} where A is a number. In the second order, our A is a matrix.

In our Fibonacci case, we can find $F_n = \left(\frac{1+\sqrt{5}}{2}\right)^n F_0$ and $F_n = \left(\frac{1-\sqrt{5}}{2}\right)^n F_0$. So we have $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda^2 \\ \lambda \end{bmatrix}$.

So one solution is $F_n = c_1 \left(\frac{1+\sqrt{5}}{2}\right)^n + c_2 \left(\frac{1-\sqrt{5}}{2}\right)^n$.

This is more general than Fibonacci. If

$$\begin{bmatrix} F_1 \\ F_1 \end{bmatrix} = A \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = A \begin{bmatrix} F_1 \\ F_0 \end{bmatrix}$$

ff (I think using the fact that $F_1 = \lambda F_0$ or something.)