Math 240 Summer 2019 Notes

Yao-Rui

Lecture notes might contain typos.

Contents

1 Lecture 1 – Introduction

1.1 Brief Overview of the Course

This course is an invitation to differential equations. What is a differential equation? Broadly speaking, it is an equation that relates functions to their derivatives. A lot of the differential equations that we study comes from physics, and our main techniques to understand these equations are linear algebra and multivariable calculus.

Ordinary Differential Equations

A standard topic in introductory physics is radioactive decay, where the associated differential equation is

$$\frac{dN}{dt} = -kN$$

for some constant k. The solution is well-known; it is

$$N(t) = e^{-kt}N_0.$$

There are a lot more physical solutions of this form. For example, the change in density as height varies is given by

$$\frac{d\rho}{dy} = -\frac{gM}{RT}\rho.$$

All these differential equations are of first degree. One can ask for the solution to higher-degree equations. For example, Hooke's Law for an isolated frictionless body can be described by

$$\frac{d^2x}{dt^2} = -\omega^2 x.$$

Here the solution is well-known too; it is

$$x(t) = A\cos(\omega t + \phi_0).$$

A large part of the course is to study these kinds of equations, called ordinary differential equations (ODEs). We will start with ODEs having constant coefficients. In fact, we generalize and study the higher dimensional version of it, for it is unreasonable for us to only work in one dimension in physics. One obvious example is modeling the oscillations of a bridge. To do this we will need to apply techniques from linear algebra, most notably the concept of eigenvalues and eigenvectors. This will be our starting point after some warmup with one-dimensional ODEs.

There are also numerous famous ODEs that are not linear. For example, vertical upward movement with resistance in liquids can be modeled by

$$\frac{dv}{dt} = \mu g\left(1 - \frac{c}{\mu g m}v^2\right)$$

and the solution to this equation is

$$v(t) = \sqrt{\frac{\mu g m}{c}}\tanh\sqrt{\frac{\mu g c}{m}}t.$$

Another example is Newton's Law of Gravitation:

$$\vec{\boldsymbol{x}}''(t) = -\frac{GM}{\|\vec{\boldsymbol{x}}(t)\|^3}\vec{\boldsymbol{x}}(t).$$

Yet another example is the predator-prey model, given by

$$\frac{dx}{dt} = \alpha x - \beta xy,$$
$$\frac{dy}{dt} = \delta xy - \gamma y.$$

We will study how to solve these kinds of ODEs, or quantitatively sketch solutions to them if an explicit solutions cannot be easily found (for example, how to analyze equilibrium points and linearize the ODEs near these points). In fact, understanding the mathematics behind Newton's Law of Gravitation is the extra credit assignment; see subsection 21.11.

Example 1.1. Let us illustrate what we mean by "solving ODEs" and "sketching solutions quantitatively" using the following system of differential equations

$$x'(t) = -y(t), \qquad y'(t) = x(t), \qquad x(t_0) = x_0, \qquad y(t_0) = y_0.$$

We will develop techniques to solve these kinds of equations; right now this example is just to let you have a feel for the kind of stuff we will do.

Let us solve this equation. Writing $\vec{\boldsymbol{x}}(t) = (x(t), y(t))$, one can rewrite our equations as

$$\vec{\boldsymbol{x}}'(t) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(t_0) = (x_0, y_0).$$

Letting A be the matrix above, the solution to this equation is

$$\vec{\boldsymbol{x}}(t) = e^{tA}\,\vec{\boldsymbol{x}}(t_0) = \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} x_0\cos(t) - y_0\sin(t) \\ x_0\sin(t) + y_0\cos(t) \end{bmatrix}.$$

In this example it is not hard to observe that $\vec{\boldsymbol{x}}(t)$ is actually a circle: writing $r^2 = x_0^2 + y_0^2$, one sees that

$$(x(t))^2 + (y(t))^2 = r^2.$$

Of course, it is not this easy to understand the behavior of most solutions to our differential equations by staring at the explicit solution in this course, so we need to develop another technique. For this we write down the vector field

$$\vec{\boldsymbol{v}}(x, y) = (-y, x)$$

associated to our system of differential equations, and observe that $\vec{\boldsymbol{v}}^{\perp}(x, y) = (x, y)$ is conservative with associated function $f(x, y) = (x^2 + y^2)/2$ satisfying

$$\nabla f(x, y) = \vec{\boldsymbol{v}}^{\perp}(x, y),$$
$$\frac{d}{dt} f(x(t), y(t)) = \nabla f(x, y) \cdot \vec{\boldsymbol{v}}(x, y) = 0.$$

Hence the unique solution $\vec{\boldsymbol{x}}(t)$ must satisfy $f(\vec{\boldsymbol{x}}(t)) = c$ for some constant c, and substituting the initial condition tells us $c = (x_0^2 + y_0^2)/2$.

It is hard to solve ODEs, and we really only have a good general theory for linear ODEs (which is the main topic for this course). In real life it is not realistic to assume that our coefficients in our ODEs will always be constant. For example, the spring constant may decrease with time as the spring weakens. We will also study ODEs with varying coefficients, paying special attention to the case of second order ODEs. We resist giving an example in this introduction.

Calculus of Variations

In the exact sciences it is important for us to know how to minimize and maximize functions subject to constraints. It is not surprising that this is related to differential equations. A standard example is to find $f : [0, 1] \longrightarrow \mathbb{R}$ such that $f(0) = 0$ and $f(1) = 1$ and such that the arc length

$$\int_0^1 \sqrt{1 + \left(\frac{df}{dx}\right)^2}\, dx$$

is minimized. We will see how the line must be the straight line from $(0, 0)$ to $(1, 1)$, agreeing with intuition. Another example is to find a frictionless surface such that a ball rolls off of it the fastest, and this is the Brachistochrone problem. In the last bit of the course we will talk about how to maximize or minimize such functions, using something called the Euler-Lagrange Equation.

Not in Math 240: Laplace Transforms

The theory of Laplace transforms is another method of solving ODEs. This is an important tool in analysis, much like the theory of Fourier transform which you may know. Unfortunately, we will not be able to study Laplace transforms and how it comes into play when solving ODEs.

An application of this theory is to derive the solution to the impulse forcing differential equation, which is the mathematical description of a harmonic oscillator being struck by a hammer at a certain time. The linear algebraic method that we have briefly illustrated in the previous example is unsuitable for this case as the striking of the hammer is modeled by the Dirac Delta function, which is not a function in the ordinary sense (rather, it is an example of what mathematicians call a *distribution*).

Math 241: Partial Differential Equations

The above equations are all examples of ODEs. However, there are evidently lots of physical equations where the variables have more than one dependencies. For example, the wave equation is given by

$$\frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 u.$$

Such equations are called partial differential equations (PDEs), and a lot of the techniques are very similar to our analysis of ODEs, in particular Sturm-Liouville Theory. Unfortunately we will not have time to study PDEs; this is left for Math 241.

1.2 Separable ODEs

Let us start the course with a warmup, which is to solve one-dimensional separable ODEs. This is an equation of the form

$$f(y)y' = g(x)$$

for continuous functions f and g. It is clear that we simply need to integrate both sides to get a general solution, i.e. to carry out

$$\int f(y)\, dy = \int g(x)\, dx.$$

Of course one has to check where the solution makes sense.

Example 1.2. Solving

$$t + xx' = 0$$

tells us

$$x(t) = \pm\sqrt{2c - t^2}, \qquad -\sqrt{2c} \leq t \leq \sqrt{2c}.$$

Giving us initial conditions will tell us how to choose the unique solution, if possible.

Example 1.3. Solving

$$x' = e^{t+x}$$

tells us

$$x(t) = -\ln(-c - e^t), \qquad t < \ln(-c)$$

Again giving us initial conditions will tell us how to choose the unique solution, if possible. Let us observe for this example that

$$\lim_{t \to \ln(-c)^-} x(t) = \infty.$$

1.3 First Order Linear ODEs

This is a differential equation of the form

$$x'(t) = p(t)x(t) + q(t), \qquad x(t_0) = x_0.$$

In order to solve for $x(t)$ one simply consider the integrating factor $e^{-\int p(t)\,dt}$. Multiplying this gives us

$$\frac{d}{dt}\left(x(t)e^{-\int p(t)\,dt}\right) = q(t)e^{-\int p(t)\,dt}.$$

Integrating tells us

$$x(t) = e^{P(t)-P(t_0)}x_0 + e^{P(t)}\int_{t_0}^{t} e^{-P(s)}q(s)\,ds,$$

where $P(t)$ is an antiderivative of $p(t)$.

Remark. The condition $x(t_0) = x_0$ is called the *initial condition* of the ODE (left ambiguous in the above examples). In general, if we are given a differential equation without conditions, then the solution we get is called the *general solution* since it depends on undetermined constants. Initial conditions allow us to pinpoint which solution we pick among the family of general solution.

Example 1.4. Consider the ODE

$$tx' = -x + t^2, \qquad x(1) = 1$$

on the interval $t > 0$. We divide throughout by t to convert it into the form above:

$$x' = -\frac{x}{t} + t.$$

The integrating factor in this case is $e^{\ln t} = t$, and we need to solve

$$\frac{d}{dt}(tx) = t^2.$$

After integration, the general solution is

$$x(t) = \frac{t^3}{3} + \frac{c}{t}.$$

If we substitute in $x(1) = 1$, the solution is

$$x(t) = \frac{t^3}{3} + \frac{2}{3t}.$$

1.4 Bernoulli ODEs

A Bernoulli ODE can be seen as a generalization of a first order linear ODE, and is a differential equation of the form

$$x'(t) = p(t)x(t) + q(t)x^n(t), \qquad n \neq 0, 1.$$

(If $n = 0, 1$ then this ODE reduces to a first order linear ODE.) To solve for $x(t)$ one considers the change of variables

$$y = x^{1-n},$$

transforming this ODE into

$$y'(t) = (1-n)p(t)y(t) + (1-n)q(t).$$

We can then solve for the general solution $y(t)$ by an integrating factor, and $x(t) = \sqrt[1-n]{y(t)}$. If an initial condition is given, then we can find the particular solution.

1.5 Ricatti ODEs

A Ricatti ODE is a differential equation of the form

$$x'(t) = p(t) + q(t)x(t) + r(t)x^2(t).$$

We can solve this ODE if we have a *particular solution* $x_1(t)$, i.e. a function that satisfies the ODE above. If this particular solution is found, we can consider the translation $u(t) = x(t) - x_1(t)$. Substituting this relation into the above ODE tells us that

$$u' - (q + 2rx_1)u - ru^2 = 0,$$

which is a Bernoulli ODE with $n = 2$. Therefore, to solve a Ricatti ODE one considers the following steps:

- Find a particular solution $x_1(t)$.
- Solve the Bernoulli ODE $u'(t) = (q(t) + 2r(t)x_1(t))u(t) + r(t)u^2(t)$.
- The general solution will be $x(t) = x_1(t) + u(t)$.

We will study some other ODEs which can only be solved after finding a particular solution.

Example 1.5. Consider the ODE

$$x' = 2t - \frac{x}{t} + \frac{x^2}{t^3}.$$

We need to find a particular solution. Since the coefficients involves only powers of t, we guess a solution of the form

$$x(t) = ct^\alpha.$$

Substituting this into the ODE tells us that $\alpha = 2$ and $c = 1, 2$, so we pick the particular solution $x_1(t) = t^2$. The associated Bernoulli ODE is

$$u' = \frac{1}{t}u + \frac{1}{t^3}u^2.$$

To solve $u(t)$ one considers the change of variables $z = u^{1-2} = u^{-1}$, transforming this ODE into

$$z' = -\frac{1}{t}z - \frac{1}{t^3}.$$

This is now a first order ODE; solving this gives

$$z(t) = \frac{ct+1}{t^2},$$

implying

$$u(t) = \frac{t^2}{ct+1}.$$

Thus the general solution to the original ODE is

$$x(t) = x_1(t) + u(t) = t^2 + \frac{t^2}{ct+1}.$$

1.6 Reduction of Order

Reduction of order is a technique where we introduce a new variable for the first derivative. This will be an important idea throughout the course. We now demonstrate this via a simple example.

Example 1.6. Consider the equation

$$tx'' - x' = t^3.$$

If we let $y = x'$, then the equation becomes

$$ty' - y = t^3.$$

This is a first order linear ODE, and its solution is

$$y(t) = \frac{t^3}{4} + \frac{c_1}{t}.$$

Integrating $y = x'$ tells us that

$$x(t) = \frac{t^4}{16} + c_1 \ln(t) + c_2.$$

2 Lecture 2 – Matrices

This is the first of three lectures giving a crash course in linear algebra. After these three lectures we will introduce additional concepts in linear algebra that we need as we go along.

2.1 Row Echelon Form

The primary goal of introducing matrices in our course is to solve a system of linear equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2, \\ \vdots \quad \vdots \quad &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

We can arrange it in terms of matrices in two ways.

The first way is to write the coefficients as a $m \times (n+1)$ array of numbers $\left[A|\,\vec{\boldsymbol{b}}\right]$, or more explicitly

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right].$$

We can then try to solve this equation via row operations, which we explain in a bit.

The second way is to write it in terms of matrix multiplication $AX = \vec{\boldsymbol{b}}$, or more explicitly

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

To make sense of the left-hand side we define matrix multiplication between an $m \times n$ matrix and an $n \times 1$ matrix in the obvious sense, such that the rows corresponds to our system of linear equations. (We define a matrix to simply be an array of numbers; we will give it a deeper meaning next time.) The general formula for matrix multiplication will be given later, and again in the next lecture.

Let us now concentrate on the first way of writing coefficients. Clearly, subtracting, scaling, and adding rows with one another corresponds to solving equations, just like how you would in high school. Recall we denoted A be the $m \times n$ array of a_{ij}'s corresponding to the array of numbers $\left[A|\,\vec{\boldsymbol{b}}\right]$.

Definition 2.1. A matrix A is in *row echelon form* (REF) if

- rows with all zeros are below any row with nonzero entries,
- the nonzero leading coefficient, or *pivot* of a row is strictly to the right of the nonzero leading coefficient of the row above it.

In addition, if every nonzero leading coefficient of A equals 1, and every column containing a 1 has zero in every other entry, then A is in *reduced row echelon form* (RREF).

Every matrix can be reduced to REF after performing a sequence of row operations. For example, the left matrix below is in REF but not the right matrix.

$$\begin{bmatrix} 2 & 5 & 8 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 2 & 5 & 8 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

The REF for both matrices above are the same, and after performing more row operations their RREF is

$$\begin{bmatrix} 1 & 5/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

There exists a systematic way of computing REF and RREF using something called Gaussian elimination, but this is quite obvious and will not be outlined here. The method is really what you think it is.

Definition 2.2. The *rank* of A is the number of leading coefficients after performing REF.

Let us return to our system of equations $\left[A|\,\vec{\boldsymbol{b}}\right]$. Such a systems always have zero, one, or infinitely many solutions, and the way to determine this is by performing row operations until A is in REF or RREF. The system has:

- no solutions if there are inconsistencies, i.e. if the last nonzero row after performing RREF is

$$\left[\begin{array}{ccc|c} 0 & \cdots & 0 & \gamma \end{array}\right], \qquad \gamma \neq 0,$$

  corresponding to $0 = 1$;
- one solution if the system is consistent and there are no free variables, i.e. rank A equals the number of columns of A;
- infinitely many solutions if the system is consistent and there are free variables, i.e. rank A is less than the number of columns of A.

Example 2.3. Let us consider the system of equations

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ -2 & -5 & 1 & 3 \\ 3 & 5 & 0 & 1 \end{array}\right], \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ -2 & -5 & 1 & 3 \\ 3 & 5 & 0 & 0 \end{array}\right].$$

After REF, we obtain

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 3 \end{array}\right], \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

There are no solutions for the first system, and for the second system

$$x_3 = t, \qquad x_2 = 5 - 3t, \qquad x_1 = -9 - 5t.$$

2.2 Determinants and Invertible Matrices

We now restrict to the case where A is an $n \times n$ matrix. The situation is the same as before: we want to solve a system of linear equations. The previous subsection tells us how to determine if this systems has solutions. We now give a criteria to determine if the system has a unique solution. To do this, we use the second way of thinking: writing the system of equations as $AX = \vec{\boldsymbol{b}}$.

Definition 2.4. The *determinant* of an $n \times n$ matrix $A = (a_{ij})$ can be recursively defined as follow.

- If $n = 1$, then $\det A = a_{11}$.
- If $n = 2$, then $\det A = a_{11}a_{22} - a_{12}a_{21}$.
- If $n > 2$, then $\det A$ can be computed in one of two ways.

– Pick a row $(a_{j1}, \ldots, a_{jn})$ of A, and let A_i be the matrix after removing the j^{th} row and i^{th} column of A. Then

$$\det A = \sum_{i=1}^{n} (-1)^{j+i} a_{ji} \det A_i.$$

– Pick a column $(a_{1j}, \ldots, a_{nj})$ of A, and let A_i be the matrix after removing the i^{th} row and j^{th} column of A. Then

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_i.$$

One can check that $\det A$ is well-defined, and in fact equals the mathematical definition

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)};$$

we will not do this here.

Example 2.5. The determinant of a 3×3 matrix

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

is $aei + bfg + cdh - (ceg - bdi - afh)$.

Example 2.6. Using the definition above,

$$\begin{aligned} \det \begin{bmatrix} 2 & -1 & 3 & 0 \\ -3 & 1 & 0 & 4 \\ -2 & 1 & 4 & 1 \\ -1 & 3 & 0 & -2 \end{bmatrix} &= 3 \begin{bmatrix} -3 & 1 & 4 \\ -2 & 1 & 1 \\ -1 & 3 & -2 \end{bmatrix} - 0 \begin{bmatrix} -2 & 1 & 0 \\ -2 & 1 & 1 \\ -1 & 3 & -2 \end{bmatrix} + 4 \begin{bmatrix} -2 & 1 & 0 \\ -3 & 1 & 4 \\ -1 & 3 & -2 \end{bmatrix} - 0 \begin{bmatrix} -2 & 1 & 0 \\ -3 & 1 & 4 \\ -2 & 1 & 1 \end{bmatrix} \\ &= 3 \cdot (-10) + 0 + 4 \cdot (-18) + 0 \\ &= -102. \end{aligned}$$

Example 2.7. The determinant of an upper triangular matrix

$$\begin{bmatrix} a_1 & * & \cdots & * \\ 0 & a_2 & \ddots & \vdots \\ 0 & 0 & \ddots & * \\ 0 & 0 & & a_n \end{bmatrix}$$

is $a_1 \cdots a_n$. The same formula holds for lower triangular matrices.

Before we state some important facts on determinants, let us define matrix multiplication; we will restate it in the next lecture. Given two $n \times n$ matrices $A = (a_{ij})$ and $B = (b_{i'j'})$, the product AB has $(i, j)^{th}$ entry

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

Example 2.8. One has

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}, \qquad AB = \begin{bmatrix} 5 & 0 \\ 12 & 0 \end{bmatrix}.$$

Definition 2.9. Let A be an $n \times n$ matrix. If there exists another $n \times n$ matrix B such that $AB = BA = I$, then A is *invertible*, and B is called the *inverse matrix* of A. We denote B by A^{-1}.

Theorem 2.10. *Let $A = (a_{ij})$ be an $n \times n$ invertible matrix. Define*

$$A_{ij} = (-1)^{i+j} \det M_{ij},$$

where M_{ij} is the $(n-1) \times (n-1)$ matrix obtained by removing the i^{th} row and j^{th} column of A. Then

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}.$$

(Note the arrangements of A_{ij} in A.) In particular, the inverse of A is unique.

Proof. Computation. □

Example 2.11. Let us consider the system $AX = \vec{\boldsymbol{b}}$ given explicitly by

$$\begin{bmatrix} 2 & -1 & 3 \\ -3 & 1 & 0 \\ -2 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}.$$

One computes $\det A = -7$, so the system has a unique solution $X = A^{-1}\,\vec{\boldsymbol{b}}$. By the theorem above

$$A^{-1} = -\frac{1}{7} \begin{bmatrix} 4 & 7 & -3 \\ 12 & 14 & -9 \\ -1 & 0 & -1 \end{bmatrix},$$

so

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -\frac{1}{7} \begin{bmatrix} 4 & 7 & -3 \\ 12 & 14 & -9 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1/7 \\ 3/7 \\ 5/7 \end{bmatrix}.$$

2.3 More Determinant Facts

Proposition 2.12. *Let A and B be $n \times n$ matrices.*

- $\det(cA) = c^n \det A$.
- $\det A^t = \det A$, *where A^t is the transpose of A obtained by writing the rows as columns.*
- $\det(AB) = \det(A)\det(B)$.
- *If A is invertible, then* $\det(A^{-1}) = (\det A)^{-1}$.
- $\det A$ *remains the same after adding or subtracting rows (or columns) with each other.*

Proof. Computation. □

Proposition 2.13. *The following are equivalent for an $n \times n$ matrix A.*

(a) *A is invertible.*
(b) $\det A \neq 0$.
(c) $\operatorname{rank} A = n$.

Proof. Statements (b) and (c) are equivalent by the last assertion in the above proposition. Statements (a) and (b) are equivalent by the third assertion and Theorem 2.10. □

Example 2.14. It is an exercise on row reduction to see that the determinant of the Vandermonde matrix

$$\begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix}$$

is

$$\prod_{1 \le i < j \le n} (a_j - a_i).$$

We have now come to the most important theorem of this lecture.

Theorem 2.15. *Consider a system of linear equations* $AX = \vec{\boldsymbol{b}}$.

- *If* $\det A \neq 0$, *then there is a unique solution for* X *given by* $X = A^{-1}\vec{\boldsymbol{b}}$.
- *If* $\det A = 0$, *then there are either zero or infinitely many solutions for* $\vec{\boldsymbol{b}}$. *There is no solution if, after performing RREF on* $\left[A|\,\vec{\boldsymbol{b}}\right]$, *the last nonzero row is of the form*

$$[\ 0 \quad \cdots \quad 0 \mid \gamma\ ], \qquad \gamma \neq 0;$$

otherwise, there are infinitely many solutions.

Proof. This is a consolidation of everything said in this lecture. □

To end this lecture we present another method to find the solution of $AX = \vec{\boldsymbol{b}}$ if A is invertible.

Theorem 2.16 (Cramer's Rule)**.** *Consider a system of linear equations* $AX = \vec{\boldsymbol{b}}$ *where* A *is an* $n \times n$ *invertible matrix. Then*

$$x_i = \frac{\det A_i}{\det A},$$

where A_i *is the matrix* A *with the* i^{th} *column replaced by* $\vec{\boldsymbol{b}}$.

Proof. Computation. □

Example 2.17. We now solve

$$\begin{bmatrix} 2 & -1 & 3 \\ -3 & 1 & 0 \\ -2 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}.$$

using Cramer's Rule. One has $\det A = -7$, and

$$\det A_1 = \det \begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 0 \\ 3 & 1 & 4 \end{bmatrix} = -1,$$

and similarly $\det A_2 = -3$ and $\det A_3 = -5$. Hence the computation agrees with Example 2.11.

3 Lecture 3 – Linear Maps

In this lecture we focus on understanding what a linear map is. In particular, we want to say that a matrix is the same as a linear map, up to change of basis. Always let F be either $\mathbb{R}$ or $\mathbb{C}$, and let F^n be the standard n-dimensional space over F.

3.1 Basis

We all know that the dimension of F^n is n. What does this mean exactly? Mathematically speaking, this means that we can pick n vectors $v_1, \dots, v_n$ such that any point of F^n can be reached by some linear combination of these vectors. Let us make this precise.

Definition 3.1. A set of vectors $v_1, \dots, v_m$ of F^n:

- is *linearly independent* if $c_1 v_1 + \cdots + c_m v_m = 0$ implies $c_1 = \cdots = c_m = 0$;
- *spans* F^n if any vector $v \in F^n$ can be written as $v = a_1 v_1 + \cdots + a_m v_m$ for some $a_1, \dots, a_m \in F$;
- is a basis if $v_1, \dots, v_m$ is linearly independent and spans F^n.

One can check linear independence and spanning as follows. Write $v_1, \dots, v_m$ as an $m \times n$ matrix, where the i^{th} row corresponds to v_i. Then perform REF (corresponding to adding, subtracting, and scaling vectors). The set:

- is linearly independent if the number of pivots equals m, and is linearly dependent otherwise;
- spans F^n if the number of pivots equals n, and does not span otherwise.

Thus, for $v_1, \dots, v_m$ to be:

- linearly independent, necessarily $m \leq n$;
- spanning, necessarily $m \geq n$;
- a basis for F^n, necessarily $m = n$.

Of course, comparing m and n is not sufficient; we need to do REF computations.

Example 3.2. The standard basis is the basis $e_1, \dots, e_n$ of F^n, where e_i is the vector with 1 in the i^{th} coordinate and 0 elsewhere.

Example 3.3. One can use REF to check that $(1, 2, 3), (1, 3, 2), (2, 1, 3)$ is a basis for F^3.

3.2 Matrix Representation of Linear Maps

Definition 3.4. A function $\varphi : F^m \longrightarrow F^n$ is a *linear map* if

$$\varphi(c\,\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}) = c\varphi(\vec{\boldsymbol{v}}) + \varphi(\vec{\boldsymbol{w}})$$

for any constant $c \in F$ and any vector $\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \in F^m$.

Example 3.5. The function $\varphi\colon F^2 \longrightarrow F$ by $\varphi(x, y) = x + 2y$ is a linear map.

Example 3.6. If A is an $n \times m$ matrix with coefficients in F, then the function $\varphi\colon F^m \longrightarrow F^n$ defined by $\varphi(\vec{\boldsymbol{v}}) = A\,\vec{\boldsymbol{v}}$ is a linear map.

Example 3.7. A rule of thumb to see if a function is linear is to check if the coordinates defining the function is a linear function without constants. A function with constants or other nontrivial functions in them are not linear. For example, the functions

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x+1 \end{array} \qquad\qquad \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^4 \\ (x,y) & \longmapsto & (e^x, y, x, xy^2 + x) \end{array}$$

are not linear maps.

Given any linear map, we can construct a matrix associated to it.

Definition 3.8. Let $\varphi : F^m \longrightarrow F^n$ be a linear map. The *standard matrix for* φ is the $n \times m$ matrix A such that $\varphi(v_i) = Av_i$ for all i.

Computing the standard matrix for φ is relatively simple. Let $e_1, \ldots, e_m$ be the standard basis for F^m. For each e_i, write

$$\varphi(e_i) = a_{1i}e_1 + \cdots + a_{ni}e_n.$$

Then the standard matrix for F^n is the matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}.$$

Example 3.9. The standard matrix associated to the linear map in Example 3.5 is

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

It is important to write down the matrix with respect to different bases as well.

Definition 3.10. Let $v_1, \ldots, v_m$ be a basis for F^m, and let $w_1, \ldots, w_n$ be a basis for F^n. The *matrix for* φ with respect to these bases is the $n \times m$ matrix $A = (a_{ij})$ such that

$$\varphi(v_j) = a_{1j}w_1 + \cdots + a_{nj}w_n$$

for all j.

Again, after fixing bases, computing the matrix for φ is relatively simple: write

$$\varphi(v_j) = a_{1j}w_1 + \cdots + a_{nj}w_n$$

for all j. Then the matrix we want is

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}.$$

Example 3.11. Let us consider the basis $(1,1), (1,0)$ for F^2 and 2 for F. Then, using the linear map in Example 3.5,

$$\varphi(1,1) = 3 = \frac{3}{2} \cdot 2,$$
$$\varphi(1,0) = 1 = \frac{1}{2} \cdot 2,$$

so the matrix of φ with respect to these bases is

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

One last thing to mention about matrix representations is the matrix for map composition.

Definition 3.12. Let $A = (a_{ij})$ and $B = (b_{i'j'})$ be two $n \times n$ matrices. Then the *product* AB is the $n \times n$ matrix with $(i,j)^{th}$ entry

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

Remark. Consider

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad BA = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

so matrices do not commute in general. In fact, matrices do not every satisfy the property $AB = AC$ implies $B = C$. An example is

$$C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad AC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Proposition 3.13. *Fix F^n, F^m, F^l, their bases, and linear maps $\varphi\colon F^n \longrightarrow F^m$ and $\psi\colon F^m \longrightarrow F^l$. If A and B corresponds to the matrices of the respective maps, then the matrix of their composition $\psi \circ \varphi\colon F^n \longrightarrow F^l$ is BA.*

Proof. Computation. □

3.3 Change of Basis

Earlier we mentioned that there are many different choice of basis for F^n. Suppose $v_1, \ldots, v_n$ and $v_1', \ldots, v_n'$ are two such bases. Can we construct an $n \times n$ matrix A such that $Av_i = v_i'$ for all i? Equivalently, we write to write down the matrix of the linear map $f : F^n \longrightarrow F^n$ defined by

$$f(a_1 v_1 + \cdots + a_n v_n) = a_1 v_1' + \cdots + a_n v_n',$$

where we fixed the basis $v_1, \ldots, v_n$ on both sides. How do we construct this matrix, called the *change of basis matrix*? We use the same method as before.

- For each i, write down v_i' as a linear combinations of the $v_1, \ldots, v_n$, i.e.

$$v_i' = a_{1i} v_1 + \cdots + a_{ni} v_n.$$

- The matrix we want is then

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

Example 3.14. To compute the change of basis matrix from $(1,2), (3,1)$ to e_1, e_2, we observe that

$$\begin{aligned} e_1 &= -\frac{1}{5} \cdot (1,2) + \frac{2}{5} \cdot (3,1), \\ e_2 &= \frac{3}{5} \cdot (1,2) - \frac{1}{5} \cdot (3,1). \end{aligned}$$

Hence the matrix we want is

$$\begin{bmatrix} -1/5 & 3/5 \\ 2/5 & -1/5 \end{bmatrix}.$$

Clearly every change of basis matrix is invertible, for we can easily construct its inverse linear map. We now have the following corollary.

Corollary 3.15. *Let $\varphi : F^n \longrightarrow F^n$ be a linear map, and let $\mathcal{B}$ and $\mathcal{B}'$ be two bases of F^n. Suppose A is the matrix of φ corresponding to the bases $\mathcal{B}$ and $\mathcal{B}$. If B is the change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$, then:*

(a) the matrix of φ corresponding to the bases $\mathcal{B}$ and $\mathcal{B}'$ is $B^{-1}A$;

(b) the matrix of φ corresponding to the bases $\mathcal{B}'$ and $\mathcal{B}'$ is $B^{-1}AB$.

Proof. Use Proposition 3.13. □

3.4 Subspaces and the Rank-Nullity Theorem

We now discuss the Rank-Nullity Theorem, which is useful for future computations. Before that, we need to defined what a subspace is.

Definition 3.16. A *subspace* of F^n is a subset W of F^n such that $cw_1+w_2 \in W$ for all $w_1, w_2 \in W$ and $c \in F$. (That is, W is closed under addition and scalar multiplication.)

Definition 3.17. Let $\varphi\colon F^n \longrightarrow F^n$ be a linear map, and let W be a subspace for F^n. We say that F^m is an *invariant subspace* if $\varphi(w) \in W$ for all $w \in W$. If A is the matrix of φ (with respect to fixed bases), this is equivalent to saying that $Aw \in W$ for all $w \in W$.

Example 3.18. Consider the two subset of F^3 defined by

$$\{(x, y, z) \in F^3 : x + 2y + z = 0\}, \qquad \{(x, y, z) \in F^3 : x + 2y + z = 1\}.$$

The first one is a subspace, but not the second one. Let us call this subspace W. Then it is clear that W is two-dimensional (two free variables), and a basis for W is $(1, 0, -1), (1, -1, 1)$.

Example 3.19 (Kernel and Image)**.** Let $\varphi\colon F^m \longrightarrow F^n$ be a linear map. Define the *kernel* and *image* of φ to be

$$\begin{aligned} \ker\varphi &= \{v \in F^m : \varphi(v) = 0\}, \\ \operatorname{im}\varphi &= \{\varphi(v) : v \in F^m\}. \end{aligned}$$

These are invariant subspaces of F^m and F^n. If we use the language of matrices, then the *kernel* and *image* of an $n \times m$ matrix A are

$$\begin{aligned} \ker A &= \{v \in F^m : Av = 0\}, \\ \operatorname{im} A &= \{Av : v \in F^m\}. \end{aligned}$$

The *nullity* $\operatorname{null} A$ and *rank* $\operatorname{rank} A$ of A is defined to be $\dim\ker A$ and $\dim\operatorname{im} A$ respectively. By observation or doing REF, $\operatorname{rank} A$ equals the number of pivots of A.

Example 3.20. By performing REF, the matrix

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

has rank 3 and nullity 2.

Theorem 3.21 (Rank-Nullity)**.** *Let* $\varphi\colon F^m \longrightarrow F^n$ *be a linear map. Then*

$$\dim\ker\varphi + \dim im\,\varphi = m.$$

In the language of matrices, if A *is an* $n\times m$ *matrix, then*

$$\operatorname{null} A + \operatorname{rank} A = m.$$

Proof. This is easy to see by doing REF on A. One can also prove it using abstract linear algebra, which we will not do here. ☐

Proposition 3.22. *Let* $\varphi\colon F^m \longrightarrow F^n$ *be a linear map. If* φ *is bijective, then* $m = n$*. Furthermore, the following are equivalent.*

(a) φ *is bijective.*
(b) $\ker\varphi = \{0\}$.
(c) φ *is injective.*
(d) φ *is surjective.*

Proof. If φ is bijective, then $\ker\varphi = \{0\}$ necessarily. By the Rank-Nullity Theorem $\dim\operatorname{im}\varphi = m$. But by bijectivity $\operatorname{im}\varphi = F^n$, so $m = n$ necessarily.

Now we show the equivalences. $(a) \Rightarrow (b)$ is clear. For $(b) \Rightarrow (c)$, if $\ker\varphi = \{0\}$, then $\varphi(v) = \varphi(w)$ implies $\varphi(v-w) = 0$, so $v - w \in \ker\varphi$ and $v - w = 0$, implying injectivity. For $(c) \Rightarrow (d)$, if φ is injective, then $\ker\varphi = \{0\}$, so $\dim\operatorname{im}\varphi = n$. But this implies $\operatorname{im}\varphi$ is a subset of F^n of the same dimension, so $\operatorname{im}\varphi = F^n$, implying surjectivity. For $(d) \Rightarrow (a)$, the Rank-Nullity Theorem tells us that $\dim\ker\varphi = 0$, so $\ker\varphi = \{0\}$ necessarily, implying injectivity, and together with surjectivity gives bijectivity. ☐

The proposition above gives the following corollary. This corollary tells us that, although matrices do not commute in general, an invertible matrix and its inverse do.

Corollary 3.23. *Let* A *and* B *be two* $n \times n$ *matrices. If* $AB = I$ *then* $BA = I$.

Proof. Since A is an invertible matrix, $\ker A = \{0\}$. Consider the matrix $I - BA$. Note that $A(I - BA) = 0$. If $I - BA$ were nonzero, then there exists a vector v such that $(I - BA)v \neq 0$. This would imply $A(I - BA)v \neq 0$ since $\ker A$ is trivial, a contradiction. ☐

3.5 An Overview of Abstract Vector Spaces

In general, one can give a notion of an abstract vector space V over F and show that every such V is isomorphic to F^n for some n. We will not dwell on the precise definition here, but we will just say the following. A *vector space* over F is a set V, together with addition and scalar multiplication satisfying some obvious axioms. A *subspace* of V is a subset W such that $cw_1 + w_2 \in W$ for every $w_1, w_2 \in W$ and $c \in F$.

In this course our vector space V will almost always be one of the following:

- a set of nice functions $\mathbb{R}^n \longrightarrow \mathbb{R}$, or
- the set of polynomials $\mathcal{P}_n$ of degree at most n with coefficients in $\mathbb{R}$.

The subspace W will almost always be the functions of V satisfying some linear ODE. (One can check that this is indeed a subspace.)

If U and V are two vector spaces, a *linear map* $\varphi\colon U \longrightarrow V$ is still defined to be a function satisfying $\varphi(cu_1 + u_2) = c\varphi(u_1) + \varphi(u_2)$ for any constant $c \in F$ and any vector $u_1, u_2 \in U$. In our course, a linear map is a map defined almost exclusively by a linear ODE. We use the same method as before to construct the matrix associated to a linear map.

Example 3.24. Let S be the space of real polynomials of the form $ax^3 + bx$, and consider the linear map $\varphi\colon S \longrightarrow \mathbb{R}^2$ defined by

$$\varphi(ax^3 + bx) = (a, a + b).$$

Pick bases x, x^3 and $(1, 0), (0, 1)$. Then

$$\begin{aligned} \varphi(x) &= (0, 1) = 0 \cdot (1, 0) + 1 \cdot (0, 1), \\ \varphi(x^3) &= (1, 1) = 1 \cdot (1, 0) + 1 \cdot (0, 1). \end{aligned}$$

Thus the matrix of φ is

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

Example 3.25. Consider the linear map $\varphi\colon \mathcal{P}_2 \longrightarrow \mathbb{R}$ defined by $\varphi(p) = p(1)$, or in other words

$$\varphi(ax^2 + bx + c) = a + b + c.$$

Since the dimensions of $\mathcal{P}_2$ and $\mathbb{R}$ are 3 and 1 respectively, the matrix of φ should be 1×3. With respect to the bases $1, x, x^2$ and 1, it is

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

Example 3.26. Let $M_2(\mathbb{R})$ be the space of 2×2 real matrices. The map $\varphi : M_2(\mathbb{R}) \longrightarrow \mathbb{R}$ defined by $\varphi(A) = \det(A)$ is not a linear map, for $\det(A + B) \neq \det(A) + \det(B)$ in general.

Example 3.27. Consider the linear map $\varphi\colon \mathcal{P}_2 \longrightarrow \mathcal{P}_2$ defined by $\varphi(p) = 2p'' + (x - 1)p$, or in other words

$$\varphi(ax^2 + bx + c) = 4a - b + (b - 2a)x + 2ax^2.$$

With respect to the basis $1, x, x^2$,

$$\begin{aligned} \varphi(1) &= 0, \\ \varphi(x) &= -1 + x, \\ \varphi(x^2) &= 4 - 2x + 2x^2. \end{aligned}$$

Thus the matrix of φ is

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

Example 3.28. Consider the vector space V of real polynomials in two variables x and y of degree at most two. This vector space is six dimensional, with basis $1, x, y, x^2, xy, y^2$. Let us consider the linear map $\varphi\colon V \longrightarrow V$ by

$$T(p(x, y)) = y\frac{\partial p}{\partial x} + x\frac{\partial p}{\partial y}.$$

A computation tells us that the matrix of φ with respect to this basis is

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

4 Lecture 4 – Eigenvalues and Eigenvectors

We introduce the most important part of linear algebra for this course, that is, eigenvalues and eigenvectors. We will see later that half of this course is about knowing how to compute these things. Throughout this lecture A is an $n \times n$ real matrix.

4.1 Eigenvalues and Eigenvectors

Definition 4.1. The *characteristic polynomial* of A is $\det(xI - A)$. An *eigenvalue* of A is a solution λ to $\det(xI - A) = 0$. The *multiplicity* of λ is the maximum positive integer m_λ such that $(x - \lambda)^{m_\lambda}$ divides $\det(A - xI)$. An *eigenvector* for an eigenvalue λ of A is a nonzero element of $\ker(A - \lambda I)$, or equivalently a nonzero solution to $A\vec{\boldsymbol{x}} = \lambda\vec{\boldsymbol{x}}$. The vector space $\ker(A - \lambda I)$ is called the *eigenspace* for λ.

Example 4.2. Consider the matrix

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

A computation tells us that

$$\det(xI - A) = (x - 1)(x - 2)^2$$

so A has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$, of multiplicities $m_{\lambda_1} = 1$ and $m_{\lambda_2} = 2$ respectively.

To compute the eigenvectors of λ_1, observe that

$$A - I = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & 0 \end{bmatrix}.$$

Since this matrix has rank 2, the rank-nullity theorem tells us that $\dim\ker(A - I) = 1$. By solving the equation $A\vec{\boldsymbol{x}} = \vec{\boldsymbol{x}}$, one sees that the eigenspace $\ker(A - I)$ is spanned by $(0, -1, 1)$.

Similarly, to compute the eigenvectors of λ_2, observe that

$$A - 2I = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & -1 \end{bmatrix}.$$

Since this matrix has rank 1, the rank-nullity theorem tells us that $\dim\ker(A - 2I) = 2$. By solving the equation $A\vec{\boldsymbol{x}} = 2\vec{\boldsymbol{x}}$, one sees that the eigenspace $\ker(A - 2I)$ is spanned by $(0, 1, 0)$ and $(-1, 0, 1)$.

Example 4.3. Consider the matrix

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

A computation tells us that

$$\det(xI - A) = (x - 1)^3$$

so A has a unique eigenvalue $\lambda = 1$ of multiplicity $m_\lambda = 3$. To compute the eigenvectors of λ, observe that

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

has rank 1. Thus $\dim\ker(A - I) = 2$. Solving the equation $A\vec{\boldsymbol{x}} = \vec{\boldsymbol{x}}$ tells us that the eigenspace $\ker(A - I)$ is spanned by $(1, 1, -1)$ and $(2, -1, 0)$.

Example 4.4. The eigenvalues of

$$\begin{bmatrix} -2 & -5 \\ 4 & 2 \end{bmatrix}$$

are $4i$ and $-4i$, and the eigenspaces are spanned by $(-1 + 2i, 2)$ and $(-1 - 2i, 2)$ respectively.

The above example illustrates a property of conjugate eigenvalues pairs.

Proposition 4.5. *If λ is an eigenvalue of A, then so is $\overline{\lambda}$. If $\vec{\boldsymbol{v}}$ is an eigenvector of λ, then $c(\vec{\boldsymbol{v}})$ is an eigenvector of $\overline{\lambda}$, where $c(\vec{\boldsymbol{v}})$ is the vector $\vec{\boldsymbol{v}}$ with all entries conjugated.*

Proof. Since $\det(xI - A)$ is a polynomial with real coefficients, complex roots must occur in pairs. If $A\vec{\boldsymbol{v}} = \lambda\vec{\boldsymbol{v}}$, then $c(A\vec{\boldsymbol{v}}) = c(\lambda\vec{\boldsymbol{v}})$. But $c(A) = A$ since A is a real matrix. □

Notice that there is an eigenvalue to every eigenvector by the following proposition.

Proposition 4.6. *There exists a nonzero solution to $A\vec{\boldsymbol{x}} = 0$ iff* $\det A = 0$.

Proof. If there is a nonzero solution, then $\det A = 0$. If $\det A = 0$, then $\operatorname{rank} A < n$ so $\ker A > 0$. □

Proposition 4.7. *Eigenvectors for distinct eigenvalues are linearly independent. More precisely, let $\lambda_1, \ldots, \lambda_n$ be pairwise distinct eigenvalues for a matrix A, and let v_i be a chosen eigenvector for λ_i. Then $v_1, \ldots, v_n$ are linearly independent.*

Proof. Let us proceed by induction. Suppose $a_1v_1 + a_2v_2 = 0$. Applying A gives us

$$a_1\lambda_1v_1 + a_2\lambda_2v_2 = 0.$$

On the other hand, multiplying throughout by λ_1 gives us

$$a_1\lambda_1v_1 + a_2\lambda_1v_2 = 0.$$

Therefore

$$a_2(\lambda_1 - \lambda_2)v_2 = 0,$$

implying $a_2 = 0$ since $\lambda_1 - \lambda_2 \neq 0$ and $v_2 \neq 0$. This tells us that $a_1v_1 = 0$, implying $a_1 = 0$.

Let us now suppose $a_1v_1 + \cdots + a_nv_n = 0$. Applying A, or multiplying by λ_n, gives us

$$a_1\lambda_1v_1 + \cdots + a_n\lambda_nv_n = 0,$$
$$a_1\lambda_nv_1 + \cdots + a_n\lambda_nv_n = 0.$$

Therefore

$$a_1(\lambda_1 - \lambda_n)v_1 + \cdots + a_{n-1}(\lambda_1 - \lambda_n)v_{n-1} = 0$$

implying $a_1 = \cdots = a_{n-1} = 0$ by induction. This implies $a_nv_n = 0$, giving $a_n = 0$ as well. □

Example 4.8. If A has n distinct eigenvalues, then the proposition tells us that A has a basis consisting of eigenvectors, with each eigenvector corresponding to a distinct eigenvalue. Example 4.4 gives an example of such a matrix.

Let us say briefly why we care about eigenvalues and eigenvectors. Let's say we want to solve the differential equation

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t)$$

with some initial condition $\vec{\boldsymbol{x}}_0$. We will see later that the unique solution is

$$\vec{\boldsymbol{x}}(t) = e^{tA}\,\vec{\boldsymbol{x}}_0,$$

where the matrix exponential will be defined in the next lecture. Thus we know the solution explicitly in principle since A and $\vec{\boldsymbol{x}}_0$ are given to us. However, the exponential factor is hard to compute directly. If A has a basis consisting of eigenvectors, the theory of eigenvectors and eigenvalues helps us to write $A = SDS^{-1}$ with D diagonal and S a change of basis matrix, whence

$$e^{tA} = Se^{tD}S^{-1}$$

and e^{tD} is easy to compute.

Definition 4.9. We say A is *diagonalizable* if it has a basis consisting of eigenvectors. Otherwise, A is *defective*.

Hence in order to check if an $n \times n$ matrix is diagonalizable we compute the eigenvalues and find a maximal set of linearly independent eigenvectors for each eigenvalue. If we get a total of n linearly independent eigenvectors among all eigenvalues, then our matrix is diagonalizable.

Proposition 4.10. *If A is diagonalizable, then $A = SDS^{-1}$ for some diagonal matrix D.*

Proof. This follows immediately as a consequence of our discussion on change of basis. □

It is important to know a method for decomposing diagonalizable matrices A into SDS^{-1}. The method is hidden in the various propositions above, and we flesh it out below.

- Compute the eigenvalues for an $n \times n$ matrix A by finding the roots of $\det(xI - A)$.
- For each eigenvalue λ, find a maximal set of linearly independent vectors for $\ker(A - \lambda I)$.
- If we get n such vectors $v_1, \dots, v_n$ in the previous step among all the eigenvalues of A, then A is diagonalizable. If v_i is associated to the eigenvalue λ_i, then

$$S = \left[\ v_1 \mid v_2 \mid \cdots \mid v_n\ \right], \qquad D = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & & \lambda_n \end{bmatrix}.$$

Example 4.11. We saw in Example 4.2 that the matrix

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

has

- eigenvalue 1 with associated eigenvector $(0, -1, 1)$, and

- eigenvalue 2 with associated eigenvectors $(0, 1, 0)$ and $(-1, 0, 1)$.

Thus, by the method outlined above,

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

The inverse matrix above can be computed using the adjunct formula, and

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

We will work out more examples of this in Lecture 6. Let us note for now that defective matrices are never diagonalizable (an example is the matrix in Example 4.3). However, we will see later that one can always write $A = SJS^{-1}$ in Jordan Canonical Form, and such that

$$e^{tA} = Se^{tJ}S^{-1},$$

with J being a special kind of upper-triangular matrix, and e^{tJ} easy to compute.

4.2 The Cayley-Hamilton Theorem

Let us end this lecture with an important theorem summarizing the concepts we have learnt in the past three lectures.

Theorem 4.12 (Cayley-Hamilton). *A satisfies the polynomial equation* $\det(xI - A) = 0$.

Proof. By viewing $A = (a_{ij})$ as a linear map $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ with respect to some basis $b_1, \ldots, b_n$, one sees that

$$\varphi(b_i) = \sum_{j=1}^{n} a_{ij} b_j.$$

Hence, replacing φ by the indeterminate x, and writing $b = (b_1, \ldots, b_n)$ one sees that

$$(xI - A)b = 0.$$

Multiplying by the adjunct of $xI - A$ on the left, one has

$$\det(xI - A)b = 0,$$

so $\det(xI - A)$ maps every b_i to 0. Hence $\det(xI - A)$ has full nullity and zero rank. □

One can check that all the matrices we have wrote down in this lecture satisfies the Cayley-Hamilton Theorem. An application of the Cayley-Hamilton Theorem is a method to compute the inverse of an invertible matrix that avoids anything to do with minors.

Corollary 4.13. *Let A be an invertible matrix, and write*

$$\det(xI - A) = x^n + a_{n-1}x^{n-1} + \cdots + a_0.$$

Then $a_0 \neq 0$, and

$$A^{-1} = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I).$$

Proof. If $a_0 = 0$ then A has zero as an eigenvalue, so $\det A = 0$ by proposition 4.6, a contradiction. Now, by Cayley-Hamilton

$$A^n + a_{n-1}A^{n-1} + \cdots + a_0 I = 0.$$

The corollary now follows by manipulation. □

Example 4.14. Let us use Cayley-Hamilton to compute the inverse of

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

The characteristic polynomial of A is

$$\det(xI - A) = x^3 - 2x^2 + 2x - 1,$$

so by the corollary above

$$A^{-1} = A^2 - 2A + 2I.$$

One can check that the computation agrees with Example 4.11.

5 Lecture 5 – Matrix Exponentials

Throughout this lecture A is an $n \times n$ real matrix. We explained last lecture why the matrix exponential is important in solving multidimensional linear first order ODEs with constant coefficients, but we have not even said what it is.

5.1 Definition and Example Computations

Definition 5.1. The *matrix exponential* of A is defined to be the convergent infinite power series

$$e^A = \sum_{j=0}^{\infty} \frac{1}{j!} A^j = I + A + \frac{1}{2!} A^2 + \cdots .$$

Certainly this definition is motivated from the Taylor series expansion of the one-dimensional exponential function about 0. A direct computation tells us that

$$\frac{d}{dt} e^{tA} = A e^{tA}, \qquad e^{tA} e^{-tA} = e^{-tA} e^{tA} = I.$$

However, it is important to note that in general

$$e^{A+B} \neq e^A e^B.$$

After expanding both sides out using the definition, one sees that the reason is because we do not have $AB = BA$.

Example 5.2. Here is an example illustrating the above non-equality. Let

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then

$$e^{A+B} = \begin{bmatrix} e & e-1 \\ 0 & 1 \end{bmatrix}, \qquad e^A e^B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} e & 1 \\ 0 & 1 \end{bmatrix}.$$

Proposition 5.3. *If A and B are two commuting $n \times n$ matrices, so that $AB = BA$, then*

$$e^{A+B} = e^A e^B.$$

Proof. Computation. □

It is very hard to compute the matrix exponential of a general square matrix directly. In principle the following proposition allows us to compute the matrix exponential of any matrix.

Proposition 5.4. *Let A be an $n \times n$ matrix.*

(a) If $A^k = 0$, then

$$e^A = \sum_{j=k}^{\infty} \frac{1}{j!} A^j = I + A + \frac{1}{2!} A^2 + \cdots + \frac{1}{(k-1)!} A^{k-1}.$$

(b) If $A = A_1 \oplus \cdots \oplus A_k$ is a block diagonal matrix, then

$$e^A = e^{A_1} \oplus \cdots \oplus e^{A_k}.$$

(c) *If*

$$A = \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda & 1 & 0 \\ 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

is a Jordan block, then

$$e^{tA} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \frac{t^3}{3!}e^{\lambda t} & \cdots & \frac{t^{n-1}}{(n-1)!}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & \frac{t}{2!}e^{\lambda t} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \frac{t^3}{3!}e^{\lambda t} \\ 0 & \cdots & 0 & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} \\ 0 & 0 & \cdots & 0 & e^{\lambda t} & te^{\lambda t} \\ 0 & 0 & 0 & \cdots & 0 & e^{\lambda t} \end{bmatrix}.$$

(d) *If* $A = SJS^{-1}$, *then* $e^{tA} = Se^{tJ}S^{-1}$.

Proof. Computation. □

Example 5.5. One observes that the matrix A below can be factored as

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

This is an example of a Jordan Canonical Form of a matrix, which we will learn how to compute in Lecture 7. Right now we are just interested to know e^{tA}, which is

$$\begin{aligned} e^{tA} &= \begin{bmatrix} -2 & 3 & 0 \\ 1 & 3 & 0 \\ 0 & -3 & 1 \end{bmatrix} \begin{bmatrix} e^t & 0 & 0 \\ 0 & e^t & te^t \\ 0 & 0 & e^t \end{bmatrix} \begin{bmatrix} -2 & 3 & 0 \\ 1 & 3 & 0 \\ 0 & -3 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} (t+1)e^t & 2te^t & 3te^t \\ te^t & (2t+1)e^t & 3te^t \\ -te^t & -2te^t & (1-3t)e^t \end{bmatrix}. \end{aligned}$$

Example 5.6. By computing out eigenvalues and eigenvectors as described in the previous lecture, we see that the matrix A below can be written as

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

Therefore

$$\begin{aligned} e^{A} &= \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} e & 0 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & e^2 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} e^2 & 0 & 0 \\ e^2 - e & e^2 & e^2 - e \\ e - e^2 & 0 & e \end{bmatrix}. \end{aligned}$$

5.2 Analytic Function on Matrices

The exponential function is an example of an analytic function. In general one can compute the result of a Jordan block after applying an analytic function on it.

Proposition 5.7. *Let f be an analytic function. If*

$$A = \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda & 1 & 0 \\ 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

is a Jordan block, then

$$f(A) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2!}f''(\lambda) & \frac{1}{3!}f'''(\lambda) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda) \\ 0 & f(\lambda) & f'(\lambda) & \frac{1}{2!}f''(\lambda) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \frac{1}{3!}f'''(\lambda) \\ 0 & \cdots & 0 & f(\lambda) & f'(\lambda) & \frac{1}{2!}f''(\lambda) \\ 0 & 0 & \cdots & 0 & f(\lambda) & f'(\lambda) \\ 0 & 0 & 0 & \cdots & 0 & f(\lambda) \end{bmatrix}.$$

Proof. Computation. □

Example 5.8. Consider the analytic function $f(x) = x^k$. Then, for $0 \le m \le k$,

$$\frac{1}{m!}f^{(m)}(x) = \frac{k(k-1)\cdots(k-m+1)}{m!}x^{k-m} = \binom{k}{m} x^{k-m}.$$

Hence, if

$$A = \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda & 1 & 0 \\ 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

is a Jordan block, then

$$A^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} & \binom{k}{3}\lambda^{k-3} & \cdots & \binom{k}{n-1}\lambda^{k-n+1} \\ 0 & \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \binom{k}{3}\lambda^{k-3} \\ 0 & \cdots & 0 & \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} \\ 0 & 0 & \cdots & 0 & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & 0 & \cdots & 0 & \lambda^k \end{bmatrix},$$

where we set $\binom{k}{l} = 0$ if $l > k$.

5.3 Nilpotent Matrices and Exponentiation

Consider an $n \times n$ matrix A with exactly one eigenvalue λ, i.e.

$$\det(A - xI) = (-1)^n (x - \lambda)^n.$$

The Cayley-Hamilton Theorem then implies $A - \lambda I = N$ for some matrix N satisfying $N^n = 0$.

Definition 5.9. A *nilpotent matrix* is a square matrix N such that $N^n = 0$ for some n.

It is easy to compute powers and exponentials of matrices with exactly one eigenvalue without using the exponential formula discussed earlier on.

Proposition 5.10. *Suppose A is an $n \times n$ matrix with exactly one eigenvalue λ. Then*

$$A^k = \sum_{j=0}^{\max\{k,n\}} \binom{k}{j} \lambda^{k-j} (A - \lambda I)^j.$$

Proof. The Cayley-Hamilton Theorem implies $(A - \lambda I)^n = 0$. Since λI and $A - \lambda I$ commutes,

$$A^k = (\lambda I + A - \lambda I)^k = \sum_{j=0}^{\max\{k,n\}} \binom{k}{j} \lambda^{k-j} (A - \lambda I)^j,$$

as desired. □

Proposition 5.11. *Suppose A is an $n \times n$ matrix with exactly one eigenvalue λ. Then*

$$e^{tA} = e^{t\lambda} \sum_{j=0}^{n-1} \frac{1}{j!} (A - \lambda_j I)^j t^j.$$

Proof. The Cayley-Hamilton Theorem implies $(A - \lambda I)^n = 0$. Since λI and $A - \lambda I$ commutes,

$$e^{tA} = e^{t\lambda I} e^{t(A - \lambda I)} = e^{t\lambda} \sum_{j=0}^{n-1} \frac{1}{j!} (A - \lambda_j I)^j t^j,$$

as desired. □

Example 5.12. Consider the matrix

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

with exactly one eigenvalue 1. One computes $(A - I)^2 = 0$, so

$$A^{2019} = I + 2019(A - I) = \begin{bmatrix} 6058 & -6057 \\ 6057 & -6056 \end{bmatrix}$$

$$e^{tA} = e^t (I + t(A - I)) = \begin{bmatrix} (3t+1)e^t & -3te^t \\ 3te^t & (-3t+1)e^t \end{bmatrix}.$$

6 Lecture 6 – First Order Constant Linear ODEs Part 1

This lecture is a review of diagonalizable matrices.

6.1 The General Solution

In the next three lectures we focus our attention to solving ODEs of the form

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{x}}_0\,.$$

This is the multidimensional version of the first order linear ODE that we saw in Lecture 1. By using the same proof over there, the solution to this differential equation has the same form.

Proposition 6.1. *The solution to the differential equation*

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{x}}_0\,.$$

is $\vec{\boldsymbol{x}}(t) = e^{(t-t_0)A}\,\vec{\boldsymbol{x}}_0$.

Proof. Since the derivative of e^{-tA} is $-Ae^{-tA}$, the above equation is equivalent to

$$\frac{d}{dt}\left(e^{-tA}\,\vec{\boldsymbol{x}}(t)\right) = 0.$$

By integrating both sides from t_0 to t and solving for the constant, we arrive at our conclusion. □

Hence the only difficulty of solving an ODE of this type is computing the exponential matrix e^{tA}. We recall how to do it via examples for diagonalizable matrices A below, and leave the general case to the next lecture.

6.2 Examples on Diagonalizable Matrices

We demonstrate the method written out in Lecture 4 via two examples. The first one should be pretty straightforward, and the second one is slightly more complicated involving complex numbers.

Example 6.2. Consider the matrix

$$A = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}.$$

Let us try to find a formula for A^k and e^A. The matrix A has

- eigenvalue 5 with eigenvector $(1, 1)$,
- eigenvalue 4 with eigenvector $(1, 2)$,

so $A = SDS^{-1}$ where

$$S = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad D = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}, \qquad S^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

This means that

$$A^k = SD^kS^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 4^k \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2\cdot 5^k - 4^k & -5^k + 4^k \\ 2\cdot 5^k - 2\cdot 4^k & -5^k + 2\cdot 4^k \end{bmatrix},$$

$$e^A = Se^DS^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} e^5 & 0 \\ 0 & e^4 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2e^5 - e^4 & -e^5 + e^4 \\ 2e^5 - 2e^4 & -e^5 + 2e^4 \end{bmatrix}.$$

Example 6.3. Consider the matrix

$$A = \begin{bmatrix} 2 & 9 & 0 & 2 \\ -1 & 2 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

Again let us try to find formulas for A^k and e^A. The matrix A has

- eigenvalue 3 with eigenvector $(19, 1, 20, 5)$,
- eigenvalue -1 with eigenvector $(3, 1, 0, -9)$,
- eigenvalue $2 + 3i$ with eigenvector $(-3i, 1, 0, 0)$,
- eigenvalue $2 - 3i$ with eigenvector $(3i, 1, 0, 0)$.

Thus $A = SDS^{-1}$ where

$$S = \begin{bmatrix} 19 & 3 & -3i & 3i \\ 1 & 1 & 1 & 1 \\ 20 & 0 & 0 & 0 \\ 5 & -9 & 0 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2+3i & 0 \\ 0 & 0 & 0 & 2-3i \end{bmatrix}.$$

We explain how to compute D^k and e^D, leaving the computation of S^{-1}, A^k, and e^A as an exercise. Note that

$$D^k = \begin{bmatrix} 3^k & 0 & 0 & 0 \\ 0 & (-1)^k & 0 & 0 \\ 0 & 0 & (2+3i)^k & 0 \\ 0 & 0 & 0 & (2-3i)^k \end{bmatrix}.$$

To compute $(2 + 3i)^k$ we use polar coordinates: observe that

$$2 + 3i = \sqrt{13}e^{i\theta}, \qquad \theta = \arctan\frac{3}{2}.$$

Hence

$$(2 + 3i)^k = (\sqrt{13})^k e^{ik\theta} = 13^{k/2}\cos(k\theta) + 13^{k/2} i \sin(k\theta).$$

Similarly

$$(2 - 3i)^k = 13^{k/2}\cos(k\theta) - 13^{k/2} i \sin(k\theta).$$

Next

$$e^D = \begin{bmatrix} e^3 & 0 & 0 & 0 \\ 0 & e^{-1} & 0 & 0 \\ 0 & 0 & e^{2+3i} & 0 \\ 0 & 0 & 0 & e^{2-3i} \end{bmatrix},$$

and one observes that

$$e^{2+3i} = e^2\cos(3) + ie^2\sin(3), \qquad e^{2-3i} = e^2\cos(3) - ie^2\sin(3).$$

6.3 Duhamel's Formula

There are many situations where we are tasked with solving an equation of the form

$$\vec{x}'(t) = A\,\vec{x}(t) + \vec{g}(t), \qquad \vec{x}(t_0) = \vec{x}_0\,.$$

instead, and we will see plenty of examples after the midterm. Fortunately we still have a simple formula to compute the solution to this equation.

Proposition 6.4. *The solution to the differential equation*

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t) + \vec{\boldsymbol{g}}(t), \qquad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{x}}_0\,.$$

is

$$\vec{\boldsymbol{x}}(t) = e^{(t-t_0)A}\,\vec{\boldsymbol{x}}_0 + e^{tA}\int_{t_0}^{t} e^{-sA}\,\vec{\boldsymbol{g}}(s)\,ds,$$

where the integration sign means to integrate each entry of $e^{-sA}\,\vec{\boldsymbol{g}}(s)$ *independently.*

Proof. Observe that the differential equation is equivalent to

$$\frac{d}{dt}\left(e^{-tA}\,\vec{\boldsymbol{x}}(t)\right) = e^{-tA}\,\vec{\boldsymbol{g}}(t),$$

and proceed as before. □

Example 6.5. Let us compute

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t) + (1,t), \qquad \vec{\boldsymbol{x}}(0) = (0,0),$$

where

$$A = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}.$$

is the matrix in Example 6.2. The solution is

$$\vec{\boldsymbol{x}}(t) = e^{tA}(0,0) + e^{tA}\int_0^t e^{-sA}(1,s)\,ds = e^{tA}\int_0^t e^{-sA}(1,s)\,ds.$$

A previous computation tells us that

$$e^{tA} = \begin{bmatrix} 2e^{5t} - e^{4t} & -e^{5t} + e^{4t} \\ 2e^{5t} - 2e^{4t} & -e^{5t} + 2e^{4t} \end{bmatrix},$$

and so

$$e^{-sA}(1,s) = \begin{bmatrix} 2e^{-5s} - e^{-4s} & -e^{-5s} + e^{-4s} \\ 2e^{-5s} - 2e^{-4s} & -e^{-5s} + 2e^{-4s} \end{bmatrix}\begin{bmatrix} 1 \\ s \end{bmatrix} = \begin{bmatrix} 2e^{-5s} - e^{-4s} - se^{-5s} + se^{-4s} \\ 2e^{-5s} - 2e^{-4s} - se^{-5s} + 2se^{-4s} \end{bmatrix}.$$

Integrating,

$$\int_0^t e^{-sA}(1,s)\,ds = \begin{bmatrix} \frac{1}{5}te^{-5t} - \frac{9}{25}e^{-5t} - \frac{1}{4}te^{-4t} + \frac{3}{16}e^{-4t} + \frac{69}{400} \\ \frac{1}{5}te^{-5t} - \frac{9}{25}e^{-5t} - \frac{1}{2}te^{-4t} + \frac{3}{8}e^{-4t} - \frac{3}{200} \end{bmatrix},$$

and $\vec{\boldsymbol{x}}(t)$ equals the product of e^{tA} and the above matrix.

7 Lecture 7 – First Order Constant Linear ODEs Part 2

In this lecture we discuss how to do a change of basis on defective matrices suitably so that one can imitate the computations done in the previous lecture. This decomposition is usually called the Jordan Canonical Form of the matrix. Throughout this lecture A will be an $n \times n$ matrix, which we view as an endomorphism $A : V \longrightarrow V$ of a complex vector space V.

7.1 Generalized Eigenvectors

Definition 7.1. Let λ be an eigenvalue of A. A *generalized eigenvector* of λ is a vector v such that $(A - \lambda I)^k v = 0$ for some positive integer k. The *generalized eigenspace* of λ is

$$E_\lambda := \{v \in V : (A - \lambda I)^k v = 0 \text{ for some positive integer } k\}.$$

The goal of this subsection is to show that V can be decomposed into a direct sum of generalized eigenspaces corresponding to the eigenvalues of A, and to show that each generalized eigenspace has dimension equal to the multiplicity of its eigenvalue.

Lemma 7.2. *Let $A : V \longrightarrow A$ be a linear map on a complex n-dimensional vector space V.*

(a) After a change of basis, A can be written as an upper triangular matrix.

(b) If $\ker A^l = \ker A^{l+1}$, then $\ker A^{l+d} = \ker A^{l+1+d}$ for any positive integer d. Also, we must have $\ker A^n = \ker A^{n+1}$.

(c) If $\operatorname{im} A^{l+1} = \operatorname{im} A^l$, then $\operatorname{im} A^{l+1-d} = \operatorname{im} A^{l-d}$ for any positive integer $d < l$. Also, we must have $\operatorname{im} A^{n+1} = \operatorname{im} A^n$.

Proof. (a) Pick an eigenvector v^λ of A corresponding to an eigenvalue λ, which exists by the fundamental theorem of algebra. After completing this to a basis $v^\lambda, v_1, \ldots, v_{n-1}$ of A, one sees that with respect to this basis

$$A = \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & * & \cdots & * \\ 0 & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{bmatrix}.$$

Now repeat this process with the bottom right $(n-1) \times (n-1)$ matrix, which corresponds to a linear map on the invariant subspace of V spanned by $v_1, \ldots, v_{n-1}$.

(b) Let $v \in \ker A^{l+1+d}$. Then $A^{l+1} A^d v = 0$, so $A^d v \in \ker A^l$. Thus $v \in \ker A^{l+1}$. If $\ker A^n \neq \ker A^{n+1}$, then $\ker A^n \subsetneq \ker A^{n+1}$. By part (b) this implies

$$\ker A \subsetneq \ker A^2 \subsetneq \cdots \ker A^n \subsetneq \ker A^{n+1},$$

so $\dim V \geq n + 1$, a contradiction.

(c) The proof is analogous to that of part (b) and will not be written here. □

Proposition 7.3. *Let λ be an eigenvalue of A with multiplicity m_λ. Then the dimension of the generalized eigenspace corresponding to λ equals m_λ.*

Proof. We proceed by induction on the dimension n of V. The proposition holds trivially if $n = 1$. For the inductive step, use the previous lemma to write

$$A = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

By definition of the characteristic polynomial the diagonal elements $\lambda_1, \dots, \lambda_n$ are exactly the eigenvalues of A, counted without multiplicity. Let $e_1, \dots, e_n$ be the basis of V with respect to this upper triangular matrix A, and let U be the invariant subspace spanned by $e_1, \dots, e_{n-1}$.

We now fix an eigenvalue λ. Then λ shows up m_λ times on the diagonal of A, and we want to show that

$$\dim \ker(A - \lambda I)^{m_\lambda} = m_\lambda.$$

We now have two cases to consider.

Case 1: $\lambda_n \neq \lambda$. By the inductive hypothesis, λ shows up exactly m_λ times on the diagonal of $A|_U$, and

$$\dim \ker(A - \lambda I)|_U^{m_\lambda} = m_\lambda.$$

We want to show that $(A - \lambda I)^{m_\lambda} v \neq 0$ for any $v \in V \setminus U$, so $\ker(A - \lambda I)|_U^{m_\lambda} = (A - \lambda I)^{m_\lambda}$, completing our claim in this case. But this is clear, for any such v can be written as $u + ce_n$ for some $u \in U$ and nonzero scalar c, and

$$(T - \lambda)^{m_\lambda}(u + ce_n) = (T - \lambda)^{m_\lambda} u + c(\lambda_n - \lambda)^{m_\lambda} e_n$$

with $c(\lambda_n - \lambda)^{m_\lambda} \neq 0$.

Case 2: $\lambda_n = \lambda$. By the inductive hypothesis, λ shows up exactly $m_\lambda - 1$ times on the diagonal of $A|_U$, and

$$\dim \ker(A - \lambda I)|_U^{m_\lambda - 1} = m_\lambda - 1.$$

By the previous lemma one has

$$\dim \ker(A - \lambda I)|_U^{m_\lambda - 1} = \dim \ker(A - \lambda I)|_U^{m_\lambda}$$

and so it suffices to show that

$$\dim \ker(A - \lambda I)^{m_\lambda} = \dim \ker(A - \lambda I)|_U^{m_\lambda} + 1.$$

By the inclusion-exclusion principle

$$\begin{aligned}\dim\left(U + \ker(A - \lambda I)^{m_\lambda}\right) &= \dim U + \dim(A - \lambda I)^{m_\lambda} - \dim\left(U \cap \ker(A - \lambda I)^{m_\lambda}\right) \\ &= (n-1) + \dim(A - \lambda I)^{m_\lambda} - \dim(A - \lambda I)|_U^{m_\lambda},\end{aligned}$$

so

$$\dim(A - \lambda I)^{m_\lambda} - \dim(A - \lambda I)|_U^{m_\lambda} \leq 1$$

as $\dim\left(U + \ker(A - \lambda I)^{m_\lambda}\right) \leq n$. Hence it suffices to find a vector of the form $u - e_n$ with $u \in U$ such that $(A - \lambda I)^{m_\lambda}(u - e_n) = 0$. By assumption

$$Ae_n = u_n + \lambda_n e_n$$

for some $u_n \in U$, so

$$(A - \lambda I)^{m_\lambda} e_n = (A - \lambda I)^{m_\lambda - 1} u_n \in \operatorname{im}(A - \lambda I)|_U^{m_\lambda - 1} = \operatorname{im}(A - \lambda I)|_U^{m_\lambda},$$

where the last equality is by the previous lemma. Thus $(A - \lambda I)^{m_\lambda} e_n = (A - \lambda I)^{m_\lambda} u$ for some $u \in U$, as desired. □

Proposition 7.4. *Let $\lambda_1, \dots, \lambda_j$ be the eigenvalues of A. Then there is an A-invariant direct sum decomposition*

$$V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_j}.$$

Proof. It is easy to see that the generalized eigenspaces intersect trivially pairwise, so the right hand side is a direct sum contained in V. By the previous proposition, $\dim E_{\lambda_l} = m_{\lambda_l}$, where m_λ is the multiplicity of the eigenvalue λ. Hence the equality is achieved, since $n = m_{\lambda_1} + \cdots + m_{\lambda_j}$. □

7.2 The Jordan Canonical Form for Nilpotent Matrices

In this subsection we consider a nilpotent matrix A, so that A has zero as the unique eigenvalue. We want to show that, after a change of basis, one can write

$$A = A_1 \oplus \cdots \oplus A_m$$

where each A_i is a block of the form

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Lemma 7.5. *Let A be a nilpotent matrix with $A^k = 0$, and let U be a nonzero subspace of V that intersects* $\ker A^{k-1}$ *trivially. Then $U + AU + \cdots + A^{k-1}U$ is A-invariant, and*

$$V = (U + AU + \cdots + A^{k-1}U) \oplus W$$

for some A-invariant subspace W.

Proof. The first claim that $U + AU + \cdots + A^{k-1}U$ is A-invariant is easy, so we omit that and prove the the second claim. We do induction on k. The case $k = 1$ is clear. For the induction hypothesis, let U' be a subspace such that

$$V = U' \oplus U \oplus \ker A^{k-1},$$

and write $U'' = U' \oplus U$. By assumption $AU'' \subset \ker A^{k-1}$. Also $AU'' \cap \ker A^{k-2} = \{0\}$, else if $Au'' \in AU''$ with $u'' \neq 0$ satisfies $A^{k-2}Au'' = 0$ then $u'' \in \ker A^{k-1}$, contradicting the direct sum decomposition of V as above.

Note that $A^{k-1} = 0$ on the A-invariant subspace $\ker A^{k-1}$. By the induction hypothesis

$$\ker A^{k-1} = (AU'' + A^2U'' + \cdots + A^{k-1}U'') \oplus W'$$

for some A-invariant subspace W'. Then Then

$$\begin{aligned} V &= U' + U + AU'' + A^2U'' + \cdots + A^{k-1}U'' + W' \\ &= (U + AU + \cdots + A^{k-1}U) + (U' + AU' + \cdots A^{k-1}U' + W'). \end{aligned}$$

Let us define

$$W = U' + AU' + \cdots + A^{k-1}U' + W',$$

which is an A-invariant subspace. It remains to show that W intersects $U + AU + \cdots + A^{k-1}U$ trivially. Suppose it does not. Then

$$u_0 + Au_1 + \cdots + A^{k-1}u_{k-1} = u'_0 + Au'_1 + \cdots + A^{k-1}u'_{k-1} + w'$$

for some $u_i \in U$ and $u'_i \in U'$ and $w' \in W'$. Applying A^{k-1} to the above equation tells us that

$$A^{k-1}(u_0 - u'_0) = 0$$

so $u_0 = u'_0$ by virtue of the direct sum decomposition of V. This implies

$$A(u_1 - u'_1) + \cdots + A^{k-1}(u_{k-1} - u'_{k-1}) = w',$$

a contradiction to the direct sum decomposition of $\ker A^{k-1}$. □

Proposition 7.6. *Let A be a nonzero nilpotent matrix. Then after a change of basis*

$$A = A_1 \oplus \cdots \oplus A_m$$

where each A_i is a block of the form

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Proof. We induct on the dimension n of V. The case $n = 1$ is trivial. For the inductive step, suppose $A^k = 0$ for some positive integer k that is chose to be the smallest. Let u be a vector chosen such that $A^{k-1}u \neq 0$, which exists otherwise we contradict the minimality of k. Consider the one-dimensional subspace U spanned by u. By the previous lemma, there is a direct sum decomposition

$$V = (U + AU + \cdots + A^{k-1}U) \oplus W$$

Hence $A = A_1 \oplus A'$, where A_1 is a linear map on $U + AU + \cdots + A^{k-1}U$, and A' is a linear map on W. Since A' is a nilpotent matrix as well and $\dim W < n$, one can apply the inductive hypothesis on A'. As for A_1, consider the vectors $A^{k-1}u, A^{k-2}u, \ldots, Au, u$, which is clearly a basis for $U + AU + \cdots + A^{k-1}U$. Then, with respect to this basis, A_1 is a block of the form as claimed in the proposition. □

7.3 The Jordan Canonical Form in General

We have now come to the main theorem in the decomposition of matrices. This result will allow us to find the exponential of any matrix easily.

Definition 7.7. Define the *Jordan block* J_k^λ to be the $k \times k$ matrix

$$J_k^\lambda = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}.$$

Theorem 7.8 (Jordan Canonical Form)**.** *Every $n \times n$ matrix A can be written as $A = SJS^{-1}$, where*

$$J = J_{k_1}^{\lambda_1} \oplus \cdots \oplus J_{k_l}^{\lambda_l}.$$

Furthermore, we have the following statements.

- *The number of Jordan blocks corresponding to an eigenvalue λ equals the number of linearly independent eigenvectors for λ.*
- *The sum of the sizes of Jordan blocks corresponding to an eigenvalue λ equals the multiplicity for λ.*
- *The number of Jordan blocks of size k equals*

$$\dim\ker(A - \lambda I)^k - \dim\ker(A - \lambda I)^{k-1}.$$

Proof. This theorem is a consolidation of all the results proven in the last two subsections. The only thing not written out explicitly is how to use the previous subsection on a general generalized eigenspace E_λ with $\lambda = 0$. For this, one considers the nilpotent map $A - \lambda I$ on E_λ. ☐

From the proof of proposition 7.6, one can extract out a method to construct the change of basis matrix S in our Jordan Canonical Form $A = SJS^{-1}$ of A. Let us restrict out attention to a generalized eigenspace E_λ.

- Compute the dimensions of the spaces $V_m = \ker(A - \lambda I)^m$.
- Let k be the largest integer such that $V_{k-1} \subsetneq V_k$. Write down a basis $v_1, \dots, v_p$ for V_{k-1}, and extend it to a basis $v_1, \dots, v_p, w_1, \dots, w_q$ for V_k.
- For each i, the vectors $(A - \lambda I)^{k-1}w_i, (A - \lambda I)^{k-2}w_i, \dots, w_i$ forms a Jordan block of size k as follow: arrange in order some consecutive rows

$$W_i = \ (A - \lambda I)^{k-1}w_i \mid (A - \lambda I)^{k-2}w_i \mid \cdots \mid w_i$$

  and associate to this the Jordan block J_k^λ of size k.
- Perform the same steps as above to the next largest integer k' such that $V_{k'-1} \subsetneq V_{k'}$, and continue doing this until no such k' exists.
- Do the same steps to all the generalized eigenspaces E_λ.
- The resulting S we want is the ordered arrangements of all the W_i's above (across all E_λ's), and the J we want is the ordered direct sum of all the Jordan blocks associated to the $W_i's$.

After using this method to compute the Jordan Canonical Form, one can use Lecture 5 to say that

$$e^{tA} = Se^{tJ}S^{-1}, \qquad A^k = SJ^kS^{-1}.$$

Example 7.9. Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Note that 1 is the only eigenvalue of A, of multiplicity 3. A computation with the rank-nullity theorem tells us that

$$\dim\ker(A - I) = \dim\ker \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 2,$$

$$\dim\ker(A - I)^2 = \dim\ker \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 3.$$

A basis for $\ker(A - I)$ is

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

and we complete this to a basis for $\ker(A - I)^2$:

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \qquad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Observe that

$$(A-I)v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

and v_2 is an eigenvector linearly independent from these two vectors. Hence one can take

$$S = \left[\ (A-I)v_3 \mid v_3 \mid v_2\ \right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix},$$

$$J = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and $A = SJS^{-1}$, where S^{-1} can be computed by the adjunct formula or by Cayley-Hamilton.

Example 7.10. Consider the matrix

$$A = \begin{bmatrix} 2 & 0 & 1 & -3 \\ 0 & 2 & 10 & 4 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

This matrix has eigenvalues 2 and 3, of multiplicities 3 and 1 respectively.

For the eigenvalue 2, observe that

$$\dim\ker(A-2I) = \dim\ker \begin{bmatrix} 0 & 0 & 1 & -3 \\ 0 & 0 & 10 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = 2,$$

$$\dim\ker(A-2I)^2 = \dim\ker \begin{bmatrix} 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = 3.$$

A basis for $\ker(A-2I)$ is

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},$$

and we complete this to a basis for $\ker(A-2I)^2$:

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Observe that

$$(A-2I)v_3 = \begin{bmatrix} 1 \\ 10 \\ 0 \\ 0 \end{bmatrix}, \qquad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

and v_2 is an eigenvector linearly independent from these two vectors.

As for the eigenvalue 3, observe that

$$w_1 = \begin{bmatrix} -3 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

is an eigenvector.

To summarize, one can take

$$S = \left[\ (A-2I)v_3 \mid v_3 \mid v_2 \mid w_1 \ \right] = \begin{bmatrix} 1 & 0 & 0 & -3 \\ 10 & 0 & 1 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

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

and $A = SJS^{-1}$, where S^{-1} can be computed by the adjunct formula or by Cayley-Hamilton.

Example 7.11. Consider the matrix

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

Note that 2 is the only eigenvalue of A, of multiplicity 3. A computation with the rank-nullity theorem tells us that

$$\dim\ker(A-2I) = \dim\ker \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} = 1,$$

$$\dim\ker(A-2I)^2 = \dim\ker \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix} = 2,$$

$$\dim\ker(A-2I)^3 = \dim\ker \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 3.$$

A basis for $\ker(A-2I)$ is

$$v_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

We complete this to a basis for $\ker(A-I)^2$:

$$v_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

followed by a basis for $\ker(A-I)^3$:

$$v_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Observe that

$$(A-I)^2 v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad (A-I)v_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Hence one can take

$$S = \left[\ (A-I)^2 v_3 \ \middle|\ (A-I)v_3 \ \middle|\ v_3\ \right] = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

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

and $A = SJS^{-1}$, where S^{-1} can be computed by the adjunct formula or by Cayley-Hamilton.

Example 7.12. Consider the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 11 & 6 & -4 & -4 \\ 22 & 15 & -8 & -9 \\ -3 & -2 & 1 & 2 \end{bmatrix}.$$

The characteristic polynomial of A is $(x+1)^2(x-1)^2$, so it has eigenvalues -1 and 1, of multiplicities 2 and 2 respectively.

For the eigenvalue -1, observe by rank-nullity that

$$\dim\ker(A+I) = \dim\ker \begin{bmatrix} 1 & 1 & 0 & 0 \\ 11 & 7 & -4 & -4 \\ 22 & 15 & -7 & -9 \\ -3 & -2 & 1 & 3 \end{bmatrix} = 1,$$

$$\dim\ker(A+I)^2 = \dim\ker \begin{bmatrix} 12 & 8 & -4 & -4 \\ 12 & 8 & -4 & -4 \\ 60 & 40 & -20 & -24 \\ -12 & -8 & 4 & 8 \end{bmatrix} = 2.$$

A basis for $\ker(A+I)$ is

$$v_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 0 \end{bmatrix}$$

and we complete this to a basis for $\ker(A+I)^2$:

$$v_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \end{bmatrix},$$

Observe that

$$(A+I)v_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}.$$

Similarly, for the eigenvalue 1, observe by rank-nullity that

$$\dim\ker(A-I) = \dim\ker \begin{bmatrix} -1 & 1 & 0 & 0 \\ 11 & 5 & -4 & -4 \\ 22 & 15 & -9 & -9 \\ -3 & -2 & 1 & 1 \end{bmatrix} = 1,$$

$$\dim\ker(A-I)^2 = \dim\ker \begin{bmatrix} 12 & 4 & -4 & -4 \\ -32 & -16 & 12 & 12 \\ -28 & -20 & 12 & 12 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 2.$$

A basis for $\ker(A-I)$ is

$$w_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

and we complete this to a basis for $\ker(A-I)^2$:

$$w_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \qquad w_2 = \begin{bmatrix} 1/4 \\ 1/4c \\ 0 \\ 1 \end{bmatrix}.$$

Observe that

$$(A-I)w_2 = \begin{bmatrix} 0 \\ 0 \\ 1/4 \\ -1/4 \end{bmatrix}, \qquad w_2 = \begin{bmatrix} 1/4 \\ 1/4 \\ 0 \\ 1 \end{bmatrix}.$$

To summarize, one can take

$$S = \left[\ (A+I)v_2 \ \middle|\ v_2 \ \middle|\ (A-I)w_2 \ \middle|\ w_2\ \right] = \begin{bmatrix} 1 & 0 & 0 & 1/4 \\ -1 & 1 & 0 & 1/4 \\ 1 & 2 & 1/4 & 0 \\ 0 & 0 & -1/4 & 1 \end{bmatrix},$$

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

and $A = SJS^{-1}$, where S^{-1} can be computed by the adjunct formula or by Cayley-Hamilton.

8 Lecture 8 – First Order Constant Linear ODEs Part 3

In this lecture we discuss two techniques to reduce computations for special kinds of first order constant linear ODEs.

8.1 An Application of Invariant Subspaces

This trick only works if the $n \times n$ matrix A in our differential equation

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{x}}_0$$

satisfies the following two conditions:

- $\vec{\boldsymbol{x}}_0$ is 0 after the k^{th} entry,
- the first k columns of A all have zero entries after the k^{th} row.

These conditions are rigged so that A has an invariant subspace consisting of the first k entries. Hence we can do the following to simplify our computations.

- Let B be the top left $k \times k$ matrix in A, and let $\boldsymbol{y}_0$ be vector consisting of the first k entries in $\vec{\boldsymbol{x}}_0$.
- Compute $e^{tB}\boldsymbol{y}_0$.
- The solution $\vec{\boldsymbol{x}}(t)$ has $e^{tB}\boldsymbol{y}_0$ as the first k entries, and 0 after the k^{th} entry.

Example 8.1. Consider $\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t)$ with initial condition $\vec{\boldsymbol{x}}(0) = (1, 1, 0, 0, 0)$, where

$$A = \begin{bmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & -4 & 6 \\ 0 & 0 & 0 & -9 & 11 \end{bmatrix}.$$

To find the solution it suffices to compute e^{tB}, where

$$B = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}.$$

The exponential matrix for Jordan blocks then tells us that

$$\vec{\boldsymbol{x}}(t) = \begin{bmatrix} e^{3t} & te^{3t} & 0 & 0 & 0 \\ 0 & e^{3t} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} (t+1)e^{3t} \\ e^{3t} \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

8.2 Buchheim's Algorithm and Some Formulas for Exponentiating Matrices

The exponential of a general square matrix can be computed using its Jordan Canonical Form, or using the so-called Buchheim's Algorithm. We discuss the general Buchheim's Algorithm here, and give simple exponentiating formulas for 2×2 and 3×3 matrices, as well as diagonalizable matrices with distinct eigenvalues. For the purposes of this course, the best strategy is to choose to use these simple formulas whenever they apply, since it is not computationally easy to find generalized eigenvectors or use Buchheim's Algorithm.

Buchheim's Algorithm

In general one computes the exponential of a matrix using Jordan Canonical Form. However, let us discuss Buchheim's Algorithm here as well, since a similar method will be used for the next lecture.

Let A be an $n \times n$ real matrix. Then the Cayley-Hamilton Theorem tells us that A satisfies its characteristic polynomial $\det(A - xI)$. Use the Fundamental Theorem of Algebra to factorize the characteristic polynomial as

$$(A - \lambda_1 I)^{n_1} \cdots (A - \lambda_k I)^{n_k} = 0.$$

Then we know that each λ_j is an eigenvalue of A, and there exists an eigenvector v_j corresponding to each λ_j.

We now want to show that the general solution to

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t)$$

can be written as

$$\vec{\boldsymbol{x}}(t) = \sum_{j=1}^{k} \sum_{i=1}^{n_j} t^{i-1} e^{\lambda_j t} C_{j,i}$$

for some constant $1 \times n$ matrices $C_{j,i}$. To do this, we need to use the fact that the space of solutions is of dimension n (for, after specifying the initial condition, one has a unique solution). With this fact, it suffices to show that the space of solutions for

$$(A - \lambda_j I)^{n_j}\,\vec{\boldsymbol{x}}(t) = \left(\frac{d}{dt} - \lambda_j\right)^{n_j} \vec{\boldsymbol{x}}(t) = 0$$

is of dimension n_j. Write our general solution as $u(t)e^{\lambda_j t} v_j$. Then

$$(A - \lambda_j I)^{n_j} u(t) e^{\lambda_j t} v_j = \frac{d^{n_j}}{dt^{n_j}}(u(t)) e^{\lambda_j t} v_j = 0,$$

so $u(t)$ is a polynomial in t of degree $n_j - 1$, completing our claim.

On the other hand, the general solution to $\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t)$ can also be written in terms of the exponential matrix as $\vec{\boldsymbol{x}}(t) = e^{tA}C$, where C is determined by the initial condition. We now describe Buchheim's Algorithm, which basically reconciles these two ways of computing $\vec{\boldsymbol{x}}(t)$.

Proposition 8.2 (Buchheim's Algorithm)**.** *Let A be an $n \times n$ matrix, and suppose A has eigenvalues λ_j of multiplicities n_j. For each eigenvalue λ_j, associate to it the numbers*

$$e^{\lambda_j t}, te^{\lambda_j t}, \ldots, t^{n_j - 1} e^{\lambda_j t}.$$

Then one can write

$$e^{tA} = \sum_{j=1}^{k} \sum_{i=1}^{n_j} t^{i-1} e^{\lambda_j t} C_{j,i}$$

for some constant $n \times n$ matrices $C_{j,i}$.

Proof. Consider the differential equation

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(0) = \vec{\boldsymbol{x}}_0\,.$$

By the arguments above one sees that

$$e^{tA}\,\vec{\boldsymbol{x}}_0 = \sum_{j=1}^{k} \sum_{i=1}^{n_j} t^{i-1} e^{\lambda_j t} C_{j,i},$$

for some constant $1 \times n$ matrices $C_{j,i}$ depending on $\vec{\boldsymbol{x}}_0$. By choosing $\vec{\boldsymbol{x}}_0$ to be one of the standard basis $e_1, \dots, e_n$, one sees that the c^{th} column of e^{tA} can be written as

$$\sum_{j=1}^{k} \sum_{i=1}^{n_j} t^{i-1} e^{\lambda_j t} C^c_{j,i}$$

for some constant $1 \times n$ matrices $C^c_{j,i}$. This completes the proof of the proposition. □

Of course, using the fact that complex eigenvalues $\lambda = \alpha \pm \beta i$ occurs in conjugate pairs, after a linear change of variables we can replace every instance of $e^{(\alpha \pm \beta i)t}$ with $e^{\alpha t} \cos \beta t$ and $e^{\alpha t} \sin \beta t$. We now demonstrate Buchheim's Algorithm on the following example.

Example 8.3. Let us compute e^{tA} for

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1/8 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}$$

A computation tells us that A has eigenvalues $\lambda_1 = 3/4$ and $\lambda_2 = 1$, both of multiplicities two. Buchheim's Algorithm then tells us that

$$e^{tA} = e^{3/4t} C_{1,1} + t e^{3/4t} C_{1,2} + e^t C_{2,1} + t e^t C_{2,2}$$

for some constant 4×4 matrices $C_{j,i}$. To evaluate these constants we solve the equations

$$\begin{aligned} I &= e^{0A} = C_{1,1} + C_{2,1} \\ A &= \left.\frac{d}{dt}\right|_{t=0} e^{tA} = \frac{3}{4} C_{1,1} + C_{1,2} + C_{2,1} + C_{2,2} \\ A^2 &= \left.\frac{d^2}{dt^2}\right|_{t=0} e^{tA} = \frac{9}{16} C_{1,1} + \frac{3}{2} C_{1,2} + C_{2,1} + 2C_{2,2} \\ A^3 &= \left.\frac{d^3}{dt^3}\right|_{t=0} e^{tA} = \frac{27}{64} C_{1,1} + \frac{27}{16} C_{1,2} + C_{2,1} + 3C_{2,2} \end{aligned}$$

to get

$$\begin{aligned} C_{1,1} &= 128A^3 - 366A^2 + 288A - 80I \\ C_{1,2} &= 16A^3 - 44A^2 + 40A - 12I \\ C_{2,1} &= -128A^3 + 366A^2 - 288A + 80I \\ C_{2,2} &= 16A^3 - 40A^2 + 33A - 9I. \end{aligned}$$

Note that the constant above can be gotten by taking the inverse matrix of the coefficients of the equations of I, A, A^2, A^3. After computing $C_{j,i}$, we see that

$$e^{tA} = \begin{bmatrix} e^t & te^t & (8t-48)e^t + (4t+48)e^{3t/4} & (16-2t)e^t + (-2t-16)e^{3t/4} \\ 0 & e^t & 8e^t + (-t-8)e^{3t/4} & -\frac{1}{2}(4e^t + (-t-4)e^{3t/4}) \\ 0 & 0 & \frac{1}{4}(t+4)e^{3t/4} & -\frac{1}{8}te^{3t/4} \\ 0 & 0 & \frac{1}{2}te^{3t/4} & -\frac{1}{4}(t-4)e^{3t/4} \end{bmatrix}.$$

One can also compute the Jordan Canonical Form of A and check that it agrees with this answer.

2×2 Matrices

Here are the exponentiation formulas for 2×2 matrices.

Proposition 8.4. *Let A be a 2×2 matrix.*

(a) If A has two distinct eigenvalues λ_1 and λ_2, then

$$e^{tA} = \frac{e^{\lambda_1 t}}{\lambda_1 - \lambda_2}(A - \lambda_2 I) + \frac{e^{\lambda_2 t}}{\lambda_2 - \lambda_1}(A - \lambda_1 I)$$

(b) If A has a unique eigenvalue λ, then

$$e^{tA} = e^{\lambda t}(I + t(A - \lambda I)).$$

Part (b) was proven in the previous subsection, and Part (a) is a special case of Theorem 8.7. Let us consider an example.

Example 8.5. If A has distinct complex eigenvalues $a + ib$ and $a - ib$, then an application of the previous proposition, together with the identity $e^{i\theta} = \cos\theta + i\sin\theta$, tells us that

$$e^{tA} = e^{at}\cos(bt)I + \frac{1}{b}e^{at}\sin(bt)(A - aI).$$

In particular, if

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

then

$$e^{tA} = e^{at}\begin{bmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt), \end{bmatrix}.$$

Note that this example and the previous proposition completely characterizes the exponential matrix of 2×2 real matrices.

3×3 Matrices

Here are the exponentiation formulas for 3×3 matrices.

Proposition 8.6. *Let A be a 3×3 matrix.*

(a) If A has three distinct eigenvalues λ_1 and λ_2, then

$$\begin{aligned} e^{tA} = & \frac{e^{\lambda_1 t}}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)}(A^2 - (\lambda_2 + \lambda_3)A + \lambda_2\lambda_3 I) \\ & + \frac{e^{\lambda_2 t}}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3)}(A^2 - (\lambda_1 + \lambda_3)A + \lambda_1\lambda_3 I) \\ & + \frac{e^{\lambda_3 t}}{(\lambda_3 - \lambda_2)(\lambda_3 - \lambda_1)}(A^2 - (\lambda_1 + \lambda_2)A + \lambda_1\lambda_2 I). \end{aligned}$$

(b) If A has two eigenvalues λ_1 and λ_2 of multiplicities one and two, then

$$e^{tA} = e^{\lambda_2 t}I + te^{\lambda_2 t}(A - \lambda_2 I) + \frac{e^{\lambda_1 t} - e^{\lambda_2 t} + (\lambda_2 - \lambda_1)te^{\lambda_2 t}}{(\lambda_2 - \lambda_1)^2}(A - \lambda_2 I)^2$$

(c) If A has a unique eigenvalue λ, then

$$e^{tA} = e^{\lambda t}\left(I + t(A-\lambda I) + \frac{t^2}{2}(A-\lambda I)^2\right).$$

Part (c) was proven in the previous subsection, and Part (a) is a special case of Theorem 8.7. We now prove part (b) by demonstrating Buchheim's Algorithm. If A has two eigenvalues λ_1 and λ_2 of multiplicities one and two, then

$$e^{tA} = e^{\lambda_1 t}B + e^{\lambda_2 t}C + te^{\lambda_2 t}D$$

for some constant 3×3 matrices B, C, D. To find these constants we solve

$$\begin{aligned} I &= B + C \\ A &= \lambda_1 B + \lambda_2 C + D \\ A^2 &= \lambda_1^2 B + \lambda_2^2 C + 2\lambda_2 D \end{aligned}$$

to get

$$\begin{aligned} B &= \frac{1}{(\lambda_2-\lambda_1)^2}(A-\lambda_2 I)^2 \\ C &= \frac{1}{(\lambda_2-\lambda_1)^2}(A-\lambda_2 I+\lambda_1-\lambda_2)(\lambda_1-A) \\ D &= \frac{1}{(\lambda_2-\lambda_1)}(A-\lambda_1)(A-\lambda_2) \end{aligned}$$

This gives us our desired formula for e^{tA} after rearranging.

Diagonalizable Matrices

For diagonalizable matrices with distinct eigenvalues we have the following formula.

Theorem 8.7 (Sylvester's Formula)**.** *Suppose A is a diagonalizable $n\times n$ matrix with distinct eigenvalues $\lambda_1,\dots,\lambda_n$. Then*

$$e^{tA} = \sum_{i=1}^{n} e^{\lambda_i t}\prod_{j\neq i}\frac{1}{\lambda_i-\lambda_j}(A-\lambda_j I).$$

Proof Sketch. Use Buchheim's Algorithm and a careful analysis with Vandermonde matrices. We omit the details here as this theorem will not be needed in this course. Of course you can still use Sylvester's Formula in this course, and the cases for $n=2$ and $n=3$ and $n=4$ can be derived easily by hand. □

9 Lecture 9 – Equilibrium Points

Up till now we have considered linear ODEs of the form $\vec{x}'(t) = A\,\vec{x}(t)$, where A is an $n \times n$ real matrix. In this lecture we will talk about nonlinear analysis around equilibrium points.

9.1 Linearization

Let $\vec{x}(t) = (x_1(t), \dots, x_n(t))$. Consider a nonlinear ODE of the form

$$\vec{x}'(t) = (f_1, \dots, f_n),$$

where f_i are functions in $x_1(t), \dots, x_n(t)$. An example of such an ODE is

$$\frac{dx}{dt} = (y-x)(1-x-y), \qquad \frac{dy}{dt} = \frac{x}{2} + xy.$$

(This ODE is nonlinear because there are terms of order 2.)

We would like to describe the solution to this ODE. The explicit methods discussed in the previous lectures will not work here since our ODE is nonlinear, and in fact we can't solve these ODEs in general. However, we can try to describe the behavior of solutions with initial conditions near equilibrium points.

Definition 9.1. An *equilibrium point* for our nonlinear ODE is a point $\vec{x}_* = (x_1, \dots, x_n)$ such that $f_i(x_1, \dots, x_n) = 0$ for all i.

If we factor in the time variable, an equilibrium point $\vec{x}_*$ is a constant solution to the nonlinear ODE after fixing a starting time t_0. Let us now write $\vec{v}(\vec{x}) = (f_1, \dots, f_n)$, i.e. the nonlinear ODE $\vec{x}'(t)$ but forgetting that each x_i depends on t. By the multivariable Taylor expansion, for every $\vec{x}$ close to $\vec{x}_*$,

$$\vec{v}(\vec{x}) \approx \vec{v}(\vec{x}_*) + [D\vec{v}(\vec{x}_*)](\vec{x} - \vec{x}_*) = [D\vec{v}(\vec{x}_*)](\vec{x} - \vec{x}_*),$$

where recall $[D\vec{v}]$ is the Jacobian matrix

$$[D\vec{v}] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix},$$

and $[D\vec{v}(\vec{x}_*)]$ is this matrix evaluated at the point $\vec{x}_*$.

Definition 9.2. The *linearization* of a nonlinear ODE at an equilibrium point x_* is the linear ODE

$$\vec{x}'(t) = [D\vec{v}(\vec{x}_*)](\vec{x}(t) - \vec{x}_*).$$

Example 9.3. Let us consider the example above:

$$\frac{dx}{dt} = (y-x)(1-x-y), \qquad \frac{dy}{dt} = \frac{x}{2} + xy.$$

The four equilibrium points are

$$(0,0), \qquad (0,1), \qquad (-1/2, -1/2), \qquad (3/2, -1/2).$$

The Jacobian matrix of this nonlinear ODE is

$$[D\vec{v}] = \begin{bmatrix} 2x-1 & -2y+1 \\ y+\frac{1}{2} & x \end{bmatrix},$$

and $[D\vec{v}(\vec{x}_*)]$ for the respective four equilibrium points are

$$\begin{bmatrix} -1 & 1 \\ \frac{1}{2} & 0 \end{bmatrix}, \qquad \begin{bmatrix} -1 & -1 \\ \frac{3}{2} & 0 \end{bmatrix}, \qquad \begin{bmatrix} -2 & 2 \\ 0 & -\frac{1}{2} \end{bmatrix}, \qquad \begin{bmatrix} 2 & 2 \\ 0 & \frac{3}{2} \end{bmatrix}.$$

9.2 Stability

We now focus our attention to two-dimensional nonlinear ODEs, i.e. when we have only two variables x_1 and x_2. In this case the Jacobian matrix have either both real eigenvalues, or complex conjugate eigenvalues (since the characteristic polynomial is of degree two).

Proposition 9.4. *Consider a two-dimensional nonlinear ODE. Let* $\vec{\boldsymbol{x}}_*$ *be an equilibrium point. After computing the eigenvalues of the Jacobian matrix* $[D_{\vec{\boldsymbol{v}}}]$*, the equilibrium point* $\vec{\boldsymbol{x}}_*$ *is a:*

- *stable node if the eigenvalues are both real and negative;*
- *unstable node if the eigenvalues are both real and positive;*
- *saddle point if the eigenvalues are both real with opposite signs;*
- *stable spiral if the eigenvalues are both complex with negative real part;*
- *unstable spiral if the eigenvalues are both complex with positive real part;*
- *stable center if the eigenvalues are both purely imaginary.*

Proof. Casework using elements from previous lectures. □

The case of zero eigenvalues is left out since analysis of it is slightly more difficult (it has both stable and unstable components). We will not dwell on this case in this course.

Example 9.5. Let us return to Example 9.3. Then one sees that:

- $(0, 0)$ is a saddle;
- $(0, 1)$ is a stable spiral;
- $(-1/2, -1/2)$ is a stable node;
- $(3/2, -1/2)$ is an unstable node.

10 Lecture 10 – Some ODEs Related to First Order ODEs

In this lecture we discuss five other kinds of ODEs that we can solve easily. The first four are closely related to first order constant linear ODEs, and the last is related to the next lecture.

10.1 Higher Order Constant Linear ODEs

Consider an ODE of the form

$$c_n y^{(n)}(t) + c_{n-1} y^{(n-1)}(t) + \cdots + c_1 y'(t) + c_0 y(t) = 0$$

with c_j constant real numbers. By introducing the vector $\vec{\boldsymbol{x}}(t) = (y^{(n-1)}(t), \ldots, y'(t), y(t))$, one observes that the above equation is equivalent to the first row of

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t)$$

for a determined $n \times n$ constant matrix A. Using the same argument in Buchheim's Algorithm, after computing $\det(\lambda I - A)$ we are led to the following method for solving this kind of equation.

- Solve the characteristic equation

$$c_n \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1 \lambda + c_0 = 0.$$

- If λ is a real solution with multiplicity k, associate to it the functions

$$e^{\lambda t}, te^{\lambda t}, \ldots, t^{k-1}e^{\lambda t}.$$

- If $\lambda = \alpha \pm \beta i$ is a pair of complex solutions with multiplicity k, associate to it the functions

$$e^{\alpha t}\cos\beta t, e^{\alpha t}\sin\beta t, te^{\alpha t}\cos\beta t, te^{\alpha t}\sin\beta t, \ldots, t^{k-1}e^{\alpha t}\cos\beta t, t^{k-1}e^{\alpha t}\sin\beta t.$$

- The general solution to the ODE above is a linear combination of the functions above, where the coefficients are determined by the initial conditions.

Of course, as was hinted in Lecture 2, any such ODE can be converted into a first order linear ODE. This is not needed in this case, and we will delay this discussion to Lecture 13 where we consider a more general kind of ODE.

Example 10.1. Hooke's Law is the differential equation

$$my''(t) = -ky(t),$$

where m is the mass of your object and k is the spring constant. By solving this equation, we see that the general solution is

$$y(t) = c_1 \cos\omega t + c_2 \sin\omega t, \qquad \omega = \sqrt{\frac{k}{m}}.$$

By defining $A = \sqrt{c_1^2 + c_2^2}$, and the angle δ such that

$$\cos\delta = \frac{c_1}{A}, \qquad \sin\delta = -\frac{c_2}{A},$$

one can also write $y(t)$ as

$$y(t) = A\cos(\omega t + \delta).$$

The constants ω, A, δ are usually called frequency, amplitude, and phase angle respectively. In Lecture 13 we will see how to solve the resonance equation, which is Hooke's Law together with some noise that tells us the unboundedness of solutions when the frequency matches external frequency.

Example 10.2. In order to solve

$$y'''(t) - 2y''(t) - 4y'(t) + 8y(t) = 0, \qquad y(0) = 0,\, y'(0) = 2,\, y''(0) = 4,$$

one computes the characteristic equation

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

telling us the general solution is

$$y(t) = c_1 e^{2t} + c_2 t e^{2t} + c_3 e^{-2t}.$$

The coefficient can be found by substituting our initial conditions into our general solution:

$$\begin{aligned} y(0) &= 0 = c_1 + c_3, \\ y'(0) &= 2 = 2c_1 + c_2 - 2c_3, \\ y''(0) &= 4 = 4c_1 + 4c_2 + 4c_3. \end{aligned}$$

Thus

$$y(t) = \frac{1}{4}e^{2t} + te^{2t} - \frac{1}{4}e^{-2t}.$$

10.2 First Order Nonconstant Linear ODEs with Time-Commuting Matrices

We have already mentioned that the general solution to $\vec{\boldsymbol{x}}'(t) = A(t)\,\vec{\boldsymbol{x}}(t)$, where $A(t)$ is a matrix function in t, is not $\vec{\boldsymbol{x}}(t) = e^{\int A(t)\,dt}C$ for some constant matrix C. This is because of the fact that A might not be a time-commuting matrix, i.e. that $A(t_1)A(t_2) \neq A(t_2)A(t_1)$. However, if this inequality is an equality, then clearly this formula still works.

Proposition 10.3. *Consider the differential equation $\vec{\boldsymbol{x}}'(t) = A(t)\,\vec{\boldsymbol{x}}(t)$. Suppose $A(t)$ is a time-commuting matrix, so $A(t_1)A(t_2) = A(t_2)A(t_1)$. Then the solution to this differential equation is*

$$\vec{\boldsymbol{x}}(t) = e^{\int A(t)\,dt}C,$$

where C determined by the initial condition. □

We will learn the general theory of such equations without the time-commuting assumption, and some methods to solve them, after the midterm.

Example 10.4. The matrices

$$A(t) = \begin{bmatrix} 1 & -\cos t \\ \cos t & 1 \end{bmatrix}, \qquad B(t) = \begin{bmatrix} 1 & \cos t \\ \cos t & 1 \end{bmatrix}, \qquad C(t) = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$$

are all time-commuting matrix.

10.3 Recursively Coupled Systems

A *recursively coupled system* is simply a system of differential equations of the form

$$x_1'(t) = f_1(x_1, \ldots, x_n), \qquad x_2'(t) = f_1(x_2, \ldots, x_n), \qquad \cdots, \qquad x_n'(t) = f_1(x_n),$$

where $x_j'(t)$ depends only on $x_j(t), x_{j+1}(t), \ldots, x_n(t)$. This system of differential equations can solved recursively, by first solving $x_n(t)$, then solving $x_{n-1}(t)$, and so on.

Example 10.5. The differential equation

$$x'(t) = y(t), \qquad y'(t) = y(t)$$

with initial conditions $x(t_0) = x_0$ and $y(t_0) = y_0$ has solution

$$x(t) = y_0 e^{t - t_0} - 1 + x_0, \qquad y(t) = y_0 e^{t - t_0}.$$

10.4 Change of Variables

We can sometimes do a change of variables to convert a hard-looking ODE with some form symmetry into another ODE that we know how to solve. There is no systematic theory for change of variables as it depends on the problem at hand, so we shall demonstrate it via two examples.

Example 10.6. A *Cauchy-Euler Equation* is a differential equation of the form

$$a_n x^n y^{(n)}(x) + a_{n-1} x^{n-1} y^{(n-1)}(x) + \cdots + a_1 x y'(x) + a_0 y(x) = 0$$

with all the a_i constants. By introducing the change of variables $x = e^u$, one gets the identities

$$\frac{dx}{du} = e^u, \qquad x^k \frac{d^k y}{dx^k} = D(D-1)\cdots(D-k+1)y(e^u),$$

where $D = d/du$. Substituting these to the Cauchy-Euler Equation, and writing $g(u) = y(e^u)$, gives us

$$a_n D(D-1)\cdots(D-n+1)g(u) + a_{n-1}D(D-1)\cdots(D-n+2)g(u) + \cdots + a_1 Dg(u) + a_0 g(u) = 0,$$

where the derivative is now with respect to u. Subsection 10.1 gave us a complete method of finding $g(u)$, and $y(x)$ can be recovered from $g(u)$ by replacing every instance of u with $\ln x$. Summarizing, we have the following method.

- Solve the characteristic equation

$$a_n \lambda(\lambda-1)\cdots(\lambda-n+1) + a_{n-1}\lambda(\lambda-1)\cdots(\lambda-n+2) + \cdots + a_1\lambda + a_0 = 0.$$

- If λ is a real solution with multiplicity k, associate to it the functions

$$x^\lambda, (\ln x)x^\lambda, \ldots, (\ln x)^{k-1}x^\lambda.$$

- If $\lambda = \alpha \pm \beta i$ is a pair of complex solutions with multiplicity k, associate to it the functions

$$x^\alpha \cos(\beta \ln x), (\ln x)x^\alpha \cos(\beta \ln x), \ldots, (\ln x)^{k-1}x^\alpha \cos(\beta \ln x),$$
$$x^\alpha \sin(\beta \ln x), (\ln x)x^\alpha \sin(\beta \ln x), \ldots, (\ln x)^{k-1}x^\alpha \sin(\beta \ln x).$$

- The general solution $y(x)$ will then be a linear combination of the functions above, where the coefficients are determined by the initial conditions.

For example, in order to find the general solution to

$$x^2 y''(x) + xy'(x) + 36y(x) = 0,$$

one solves $\lambda(\lambda-1) + \lambda + 36 = 0$ to get $\lambda = \pm 6i$. Then the general solution will be

$$y(x) = c_1 \cos(6\ln x) + c_2 \sin(6 \ln x).$$

Example 10.7. Sometimes one can decouple a system to change it into a recursively coupled one. Consider the differential equation

$$x'(t) = y(t) + x(t)(1 - x^2(t) - y^2(t)),$$
$$y'(t) = -x(t) + y(t)(1 - x^2(t) - y^2(t)),$$

with $x(0) = x_0$ and $y(0) = y_0$. Let us assume for simplicity that $x_0 > 0$ and $y_0 > 0$, and we search for solutions starting at (x_0, y_0), i.e. with $t \geq 0$. By writing $x = r\cos\theta$ and $y = r\sin\theta$, one gets

$$\begin{aligned} r'(t) &= \frac{d}{dt}\sqrt{x^2(t) + y^2(t)} \\ &= \frac{x(t)x'(t)}{\sqrt{x^2(t) + y^2(t)}} + \frac{y(t)y'(t)}{\sqrt{x^2(t) + y^2(t)}} \\ &= r(t)(1 - r^2(t)). \end{aligned}$$

Also $\theta'(t) = -1$ by comparing the formula for $x'(t)$ above with

$$\begin{aligned} x'(t) &= \frac{d}{dt} r(t)\cos\theta(t) \\ &= r'(t)\cos\theta(t) - r(t)\theta'(t)\sin\theta(t) \\ &= r(t)(1 - r^2(t))\cos\theta(t) - r(t)\theta'(t)\sin\theta(t) \\ &= x(t)(1 - x^2(t) - y^2(t)) - \theta'(t)y(t). \end{aligned}$$

Hence to solve for $x(t) = r(t)\cos\theta(t)$ and $y(t) = r(t)\sin\theta(t)$ one just needs to solve for

$$r'(t) = r(t) - r^3(t), \qquad \theta'(t) = -1,$$

with $r(0) = \sqrt{x_0^2 + y_0^2}$ and $\theta(0) = \arctan(x_0 y_0^{-1})$. The first ODE is a Bernoulli equation. We learned from Lecture 1 that in order to solve this one substitutes $u(t) = r^{-2}(t)$ to get

$$u'(t) = -2u(t) + 2, \qquad u(0) = u_0 = \frac{1}{r_0^2}.$$

whence $u(t) = 1 - e^{-2t}(1 - u_0)$. This implies

$$r(t) = \frac{r_0 e^t}{\sqrt{r_0^2 e^{2t} - (r_0^2 - 1)}}$$

which is well-defined if we assume $t \geq 0$. The second ODE is easy; it has as solution

$$\theta(t) = \theta_0 - t.$$

Therefore by using angle formulas

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \frac{e^t}{\sqrt{1 - (x_0^2 + y_0^2) + (x_0^2 + y_0^2)e^{2t}}} \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}.$$

Notice that, if we write $\vec{\boldsymbol{x}}(t) = (x(t), y(t))$ and $\vec{\boldsymbol{x}}_0 = (x_0, y_0)$, then

$$\lim_{t\to\infty} \vec{\boldsymbol{x}}(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \frac{\vec{\boldsymbol{x}}_0}{\| \vec{\boldsymbol{x}}_0 \|}.$$

Hence any solution to our ODE stabilizes eventually, and circles around the unit circle centered at $(0, 0)$ with period 2π. The point $(0, 0)$ is an *equilibrium point*, and the unit circle centered at $(0, 0)$ is a *limit cycle*. In general understanding the behavior of ODEs is hard, and is an active area of research. We will learn some criteria to determine the stability of equilibrium points in the next lecture.

10.5 Hamiltonian ODEs

It is hard to solve a general 2-dimensional ODE. For a lot of applications it is more important to know the solution curve of an ODE, and not the explicit solution for it. In this subsection we study the case of Hamiltonian ODEs, leaving the more general situation for the next lecture. The reason why Hamiltonian ODEs are easier to study is two-fold:

- it gives rise to a conservative vector field, from which we can write down the solution curves using techniques from Calculus I and II, and
- the corresponding solution curve in the xyz-plane (ignoring parameter t) is a level curve.

This is particularly important in physical applications as it gives us a quantity that is conserved over time.

Definition 10.8. A system of ODEs of the form

$$x'(t) = f(x(t), y(t)), \qquad y'(t) = g(x(t), y(t))$$

is *Hamiltonian* if

$$\frac{\partial}{\partial x} f(x, y) + \frac{\partial}{\partial y} g(x, y) = 0.$$

Proposition 10.9. *Let $\vec{x}(t) = (x(t), y(t))$ be a solution to the Hamiltonian ODE above. Then the graph $H(x, y)$ traced out by the solution is a level curve. It can be computed by*

$$H(x, y) = -\int g(x, y)\, dx + \alpha(y) \qquad \text{or} \qquad H(x, y) = \int f(x, y)\, dy + \beta(x),$$

where $\alpha(y)$ and $\beta(x)$ can be found by equating the two ways of finding $H(x, y)$.

Proof. Consider the vector field $\vec{v}(x, y) = (f(x, y), g(x, y))$. Notice that the perpendicular vector field $\vec{v}^{\perp}(x, y) = (-g(x, y), f(x, y))$ satisfies

$$-\frac{\partial}{\partial y} g(x, y) = \frac{\partial}{\partial x} f(x, y)$$

by assumption, so we have our claim for computing $H(x, y)$. Also

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\frac{dx}{dt} + \frac{\partial H}{\partial y}\frac{dy}{dt} = -gf + fg = 0,$$

so $H(x, y)$ is a level curve. $\square$

Example 10.10. Consider

$$x'(t) = -2x^4(t)y^3(t) - y(t), \qquad y'(t) = 2x^3(t)y^4(t) + x(t),$$

with initial conditions $x(t_0) = 0$ and $y(t_0) = -2$. In this case

$$\frac{\partial}{\partial x}(-2x^4y^3 + y) = -8x^3y^3, \qquad \frac{\partial}{\partial y}(2x^3y^4 + x) = 8x^3y^3,$$

and after some computation

$$H(x, y) = \frac{-x^4y^4 - y^2 - x^2}{2}.$$

To find the level of $H(x, y)$, one substitutes the initial conditions to see that $H(0, -2) = -2$. Hence the solution $(x(t), y(t))$ to our ODE satisfies $H(x, y) = -2$, or

$$y^2 + x^4y^4 + x^2 = 4,$$

and this graph looks like a "deformed circle".

11 Lecture 11 – Midterm Review

11.1 Midterm Review

A review of lectures 1 to 9 was given during this lecture. You should be comfortable with the topics below; this list is also repeated in the next lecture.

One-Dimensional ODEs

- Separable and First Order Linear ODEs
- Bernoulli and Ricatti ODEs

Matrices

- Row Echelon Form
- Compute inverses using cofactor expansion or the Cayley-Hamilton Theorem
- Rank-Nullity Theorem
- Eigenvalues and generalized eigenvectors
- Jordan Canonical Form
- Compute powers and exponentials using Jordan Canonical Form or Buchheim's Algorithm
- Exponentiation formulas for nilpotent, diagonalizable, and small matrices

First Order Constant Linear ODEs

- Duhamel's Formula
- Reduction of order
- Linearization and stability

12 Lecture 12 – Midterm

12.1 Midterm

The midterm was given during this lecture; see subsection 21.6. Below is a non-exhaustive list of things that you should know for the midterm.

One-Dimensional ODEs

- Separable and First Order Linear ODEs
- Bernoulli and Ricatti ODEs

Matrices

- Row Echelon Form
- Compute inverses using cofactor expansion or the Cayley-Hamilton Theorem
- Rank-Nullity Theorem
- Eigenvalues and generalized eigenvectors
- Jordan Canonical Form
- Compute powers and exponentials using Jordan Canonical Form or Buchheim's Algorithm
- Exponentiation formulas for nilpotent, diagonalizable, and small matrices

First Order Constant Linear ODEs

- Duhamel's Formula
- Reduction of order
- Linearization and stability

13 Lecture 13 – Some General Theory for Linear ODEs

In this lecture we will sketch the general theory of linear ODEs with examples. Most of the proofs will be omitted as they are not essential to the course.

13.1 The Space of Solutions

Let $\vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}, t)$ be a continuous function from $\mathbb{R}^n \times [a, b]$ to $\mathbb{R}^n$ satisfying the *Lipschitz condition*:

$$\| \vec{\boldsymbol{v}}(\vec{\boldsymbol{y}}, t) - \vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}, t)\| \leq L\| \vec{\boldsymbol{y}} - \vec{\boldsymbol{x}} \|$$

for some fixed L, all $\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}} \in \mathbb{R}^n$, and all $t \in [a, b]$.

Example 13.1. An example of such a continuous function is a system of constant coefficient linear ODEs $\vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}, t) = A\,\vec{\boldsymbol{x}}$. Here we can take L to be the sum of squares of all the entries of A.

The following theorem gives the foundation for everything we do in this course.

Theorem 13.2 (Picard-Lindelöf)**.** *Fix $t_0 \in (a, b)$ and $\vec{\boldsymbol{x}}_0 \in \mathbb{R}^n$. Then there is a unique solution to*

$$\vec{\boldsymbol{x}}'(t) = \vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}, t), \qquad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{x}}_0\,.$$

This solution is continuous on $[a, b]$ and differentiable on (a, b). In particular, the space of general solutions to $\vec{\boldsymbol{x}}'(t) = \vec{\boldsymbol{v}}(\vec{\boldsymbol{x}}, t)$ has dimension n.

Proof. Omitted as it is beyond the scope of the course. □

Among other things, this theorem implies that there are n linearly independent solutions to an ODE of the form $\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t)$, or in particular an ODE of the form

$$c_n y^{(n)}(t) + c_{n-1} y^{(n-1)}(t) + \cdots + c_1 y'(t) + c_0 y(t) = 0.$$

13.2 Flows and the General Duhamel's Formula

We now specialize to the case $\vec{\boldsymbol{x}}'(t) = A(t)\,\vec{\boldsymbol{x}}(t)$, where $A(t)$ is a continuous $n \times n$ real matrix on the interval (a, b).

Theorem 13.3. *Consider*

$$\vec{\boldsymbol{x}}'(t) = A(t)\,\vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{x}}_0\,.$$

Suppose $x_1(t), \ldots, x_n(t)$ are n solutions such that the matrix

$$\left[\ x_1(t_0) \ \middle|\ x_2(t_0) \ \middle|\ \cdots \ \middle|\ x_n(t_0)\ \right]$$

is invertible. Define

$$M(t, s) = \left[\ x_1(t) \ \middle|\ x_2(t) \ \middle|\ \cdots \ \middle|\ x_n(t)\ \right] \left[\ x_1(s) \ \middle|\ x_2(s) \ \middle|\ \cdots \ \middle|\ x_n(s)\ \right]^{-1}.$$

Then $M(t, t_0)$ is invertible for all t, and the unique solution to the ODE is

$$\vec{\boldsymbol{x}}(t) = M(t, t_0)\,\vec{\boldsymbol{x}}_0\,.$$

Proof. Straightforward computation. □

Corollary 13.4. *Consider*

$$\vec{\boldsymbol{x}}'(t) = A(t)\,\vec{\boldsymbol{x}}(t) + \vec{\boldsymbol{b}}(t), \qquad \vec{\boldsymbol{x}}(t_0) = \vec{\boldsymbol{x}}_0\,.$$

If we let $M(t,s)$ be as in the theorem above, then the unique solution to this ODE is

$$\vec{\boldsymbol{x}}(t) = M(t,t_0)\,\vec{\boldsymbol{x}}_0 + \int_{t_0}^{t} M(t,s)\,\vec{\boldsymbol{b}}(s)\,ds.$$

Proof. Observe by definition of $M(t,s)$ that

$$M(s,t) = M(t,s)^{-1}, \qquad \frac{d}{dt}M(t,s) = AM(t,s), \qquad M(t,t) = 1.$$

Now, by the chain rule, for any invertible matrix T

$$0 = \frac{d}{dt}(1) = \frac{d}{dt}(TT^{-1}) = \frac{dT}{dt}T^{-1} + T\frac{d}{dt}(T^{-1}),$$

and so

$$\frac{d}{dt}(T^{-1}) = -T^{-1}\frac{dT}{dt}T^{-1}.$$

With this, use the chain rule to see that

$$\frac{d}{dt}(M(t_0,t)\,\vec{\boldsymbol{x}}(t)) = \frac{d}{dt}(M(t,t_0)^{-1}\,\vec{\boldsymbol{x}}(t)) = M(t_0,t)(\vec{\boldsymbol{x}}'(t) - A(t)\,\vec{\boldsymbol{x}}(t)).$$

This tells us that

$$\frac{d}{dt}(M(t_0,t)\,\vec{\boldsymbol{x}}(t)) = M(t_0,t)\,\vec{\boldsymbol{b}}(t).$$

Integrate both sides from t_0 to t to see that

$$\vec{\boldsymbol{x}}(t) = M(t,t_0)\,\vec{\boldsymbol{x}}_0 + M(t,t_0)\int_{t_0}^{t} M(t_0,s)\,\vec{\boldsymbol{b}}(s)\,ds$$

Since $M(t,t_0)M(t_0,s) = M(t,s)$, this gives us what we want. □

This theorem is only useful in practice if we can find n linearly independent solutions to our ODE. Unfortunately, this is the only part of the course that does not have an explicit method. In general it is very difficult, and we will learn more about finding such solutions in the next lecture for the case of second order ODEs. One should now review the types of ODEs where we have methods to find solutions, such as the ones in Lecture 10. We simplify some of them here for second-order ODEs, since this is what we are mostly dealing with.

Example 13.5. A constant-coefficient ODE of the form

$$x''(t) + bx'(t) + cx(t) = 0$$

can be solved by considering the characteristic equation

$$\lambda^2 + b\lambda + c = 0.$$

The solution to this equation is, of course,

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4c}}{2}.$$

We now have three cases:

- if $\sqrt{b^2-4c} > 0$, then the solution is

$$x(t) = c_1 e^{\frac{-b+\sqrt{b^2-4c}}{2}t} + c_2 e^{\frac{-b-\sqrt{b^2-4c}}{2})t}.$$

- if $\sqrt{b^2-4c} = 0$, then the solution is

$$x(t) = c_1 e^{-\frac{b}{2}t} + c_2 t e^{-\frac{b}{2}t}.$$

- if $\sqrt{b^2-4c} < 0$, then the solution is

$$x(t) = c_1 e^{-\frac{b}{2}t} \cos(\frac{\sqrt{b^2-4c}}{2}t) + c_1 e^{-\frac{b}{2}t} \sin(\frac{\sqrt{b^2-4c}}{2}t).$$

Example 13.6. An ODE of the form

$$t^2 x''(t) + btx'(t) + cx(t) = 0$$

with b, c constants can be solved by considering $x(t) = t^\alpha$. If we substitute this in, then we get

$$\alpha^2 + (b-1)\alpha + c = 0.$$

Thus

$$\alpha = \frac{-(b-1) \pm \sqrt{(b-1)^2 - 4c}}{2},$$

and we have three cases:

- if $\sqrt{(b-1)^2-4c} > 0$, then the solution is

$$x(t) = c_1 t^{\frac{-(b-1)\pm\sqrt{(b-1)^2-4c}}{2}} + c_2 t^{\frac{-(b-1)\pm\sqrt{(b-1)^2-4c}}{2}}.$$

- if $\sqrt{(b-1)^2-4c} = 0$, then the solution is

$$x(t) = c_1 t^{-\frac{(b-1)}{2}} + c_1 (\ln t) t^{-\frac{(b-1)}{2}}.$$

- if $\sqrt{(b-1)^2-4c} < 0$, then the solution is

$$x(t) = c_1 t^{-\frac{(b-1)}{2}} \cos(\frac{\sqrt{(b-1)^2-4c}}{2} \ln x) + c_1 t^{-\frac{(b-1)}{2}} \sin(\frac{\sqrt{(b-1)^2-4c}}{2} \ln x).$$

14 Lecture 14 – Application to Second Order ODEs Part 1

This lecture is a series of carefully worked-out examples based on the previous lecture, applied to the case of second order linear ODEs. After dividing by the leading coefficient, we consider ODEs of the form

$$y''(t) + p(t)y'(t) + q(t)y(t) = r(t).$$

The idea to solve this equation is to find two linearly independent solutions $y_1(t)$ and $y_2(t)$ to the homogenized equation

$$y''(t) + p(t)y'(t) + q(t)y(t) = 0,$$

and find a particular solution $y_p(t)$ to the original ODE. Then the general solution will be

$$y(t) = c_1y_1(t) + c_2y_2(t) + y_p(t).$$

There is no general recipe to find $y_1(t)$, but we can give recipes for $y_2(t)$ and $y_p(t)$ after finding $y_1(t)$. Of course, one can also guess for $y_2(t)$ and $y_p(t)$ if the ODE is simple enough.

14.1 Finding Solutions

As we have said a few times, finding a solution to

$$y''(t) + p(t)y'(t) + q(t)y(t) = 0$$

is the only part of the course without a good recipe. In the previous lecture (see also Lecture 10 for the general case) we considered:

- the characteristic polynomial method when $p(t)$ and $q(t)$ are constants;
- Cauchy-Euler when $p(t) = bt^{-1}$ and $q(t) = ct^{-2}$.

If both $p(t)$ and $q(t)$ are trigonometric functions, we will try to guess a trigonometric solution. Another example where one can try to guess for a solution is an ODE where both $p(t)$ and $q(t)$ are polynomials in t. For this kind of equations, we guess a solution of the form $h(t)$ or $e^{h(t)}$, where $h(t)$ is another polynomial in t.

Example 14.1. A solution to $(1-t^2)x''(t) - 2tx'(t) + 2x(t) = 0$ is $x(t) = t$.

Example 14.2. A solution to $y'' - 4ty' + (4t^2 - 2)y = 0$ is $x(t) = e^{t^2}$.

Although we will see a recipe for finding a particular solution to

$$y''(t) + p(t)y'(t) + q(t)y(t) = r(t),$$

we can sometimes guess for the particular solution if the ODE is easy enough. In the following examples we will take note of the following easy proposition.

Proposition 14.3. *For $i = 1, 2$, if $y_{p,i}(t)$ is a particular solution to*

$$y''(t) + p(t)y'(t) + q(t)y(t) = r_i(t),$$

then $y_{p,1}(t) + y_{p,2}(t)$ is a particular solution to

$$y''(t) + p(t)y'(t) + q(t)y(t) = r_1(t) + r_2(t).$$

Proof. Clear. □

The rule of thumb for guessing particular solutions are as follow:
• if $r(t)$ has an exponential function, put an exponential function in your guess,
• if $r(t)$ has a sin or a cos function, put a sum of sin and cos in your guess,,
• if $r(t)$ has a polynomial function, put a polynomial function of the same degree in your guess,,
• if $r(t)$ is a sum of functions, try to use the proposition above.

Example 14.4. A solution to $y'' + y = t$ is $y(t) = t$.

Example 14.5. A solution to $y'' - 5y' + 6y = e^t$ is $y(t) = e^t/2$.

Example 14.6. A solution to $y'' - 4y' - 12y = te^t$ is $y(t) = -\frac{1}{36}(3t+1)e^{4t}$.

Example 14.7. A solution to $y'' - 4y' - 12y = 3e^{5t} + \sin(2t)$ is $y(t) = -\frac{3}{7}e^{5t} + \frac{1}{40}\cos(2t) - \frac{1}{20}\sin(2t)$.

14.2 Duhamel's Formula in This Case

The following theorem is Duhamel's Formula for

$$y''(t) + p(t)y'(t) + q(t)y(t) = r(t).$$

After finding one solution to the equation $y''(t)+p(t)y'(t)+q(t)y(t) = 0$, it tells us how to completely solve this kind of differential equation.

Theorem 14.8 (Variation of Parameters)**.** *Consider the differential solution*

$$y''(t) + p(t)y'(t) + q(t)y(t) = r(t).$$

(a) Let $y_1(t)$ be a solution to the equation $y''(t) + p(t)y'(t) + q(t)y(t) = 0$. Then the general solution to this differential equation is $y_0(t) = c_1 y_1(t) + c_2 y_2(t)$, where

$$y_2(t) = y_1(t) \int_{t_0}^{t} \frac{1}{y_1(t)^2} e^{-P(t)}\, dt$$

and $P(t)$ is an antiderivative of $p(t)$.
(b) The general solution to $y''(t) + p(t)y'(t) + q(t)y(t) = r(t)$ is

$$y(t) = y_0(t) + y_p(t),$$

where $y_0(t) = c_1 y_1(t) + c_2 y_2(t)$ is the solution found in part (a), and

$$y_p(t) = \int_{t_0}^{t} \frac{y_1(s)y_2(t) - y_2(s)y_1(t)}{y_1(s)y_2'(s) - y_2(s)y_1'(s)} r(s)\, ds.$$

(c) If we impose initial conditions $y(t_0) = \alpha$ and $y'(t_0) = \beta$, then the coefficients c_1 and c_2 are uniquely determined by

$$c_1 = \frac{\alpha y_2'(t_0) - \beta y_2(t_0)}{y_1(t_0)y_2'(t_0) - y_2(t_0)y_1'(t_0)}, \qquad c_2 = \frac{-\alpha y_1'(t_0) + \beta y_1(t_0)}{y_1(t_0)y_2'(t_0) - y_2(t_0)y_1'(t_0)}.$$

Proof. We know the space of solutions must be 2-dimensional from the previous lecture. The rest follows by computation. □

As we have seen, in practice one does not need to use the above proposition in its entirety; one can sometimes guess for both $y_1(t)$ and $y_2(t)$, and even for $y_p(t)$ if it is simple enough.

Example 14.9. Consider the ODE

$$y''(t) + t^{-1}y'(t) - t^{-2}y(t) = t^2, \qquad y(1) = 1, \qquad y'(1) = -1.$$

Two linearly independent solutions to the homogenized equation are $y_1(t) = t$ and $y_2(t) = t^{-1}$. Plugging this into the variation of parameters formula,

$$y_p(t) = \frac{2t^5 - 5t^2 + 3}{30t}.$$

If we substitute in the initial conditions (we don't have to use the previous theorem here), the solution we seek is

$$y(t) = t^{-1} + \frac{2t^5 - 5t^2 + 3}{30t}.$$

Of course, we could have solved this by a Cauchy-Euler computation.

Example 14.10. Consider the ODE

$$2y'' + 18y = 6\tan(3t).$$

A computation tells us that

$$y(t) = c_1 \cos(3t) + c_2 \sin(3t) - \frac{1}{3}\cos(3t) \ln\left|\frac{1+\sin(3t)}{1-\sin(3t)}\right|$$

Example 14.11. Consider the ODE

$$x''(t) = -kx(t) + f(t)$$

where k is a positive constant. From the previous lecture we know that two homogeneous solutions are $x_1(t) = \cos(\sqrt{k}t)$ and $x_2(t) = \sin(\sqrt{k}t)$. By the variation of parameters formula, a particular solution is

$$x_p(t) = \frac{1}{\sqrt{k}} \int_{t_0}^{t} \sin(\sqrt{k}(t-s)) f(s)\, ds.$$

15 Lecture 15 – Application to Second Order ODEs Part 2

In this lecture we discuss the linear algebra required to understand an important kind of second order ODE called the mechanical oscillation equation. This is a higher-dimensional analog of Hooke's Law for mechanical spring.

15.1 Inner Products and Orthogonality

Again let F be either $\mathbb{R}$ or $\mathbb{C}$. Recall that the inner product on F^n is defined by

$$\langle \vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \rangle = \overline{v_1} w_1 + \cdots + \overline{v_n} w_n.$$

(If F is the field of real numbers, then $\langle \vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \rangle = v_1 w_1 + \cdots + v_n w_n$.) If A is an $n \times n$ matrix with coefficients in F, then it is not hard to see that

$$\langle \vec{\boldsymbol{x}}, A\,\vec{\boldsymbol{x}} \rangle = \vec{\boldsymbol{x}}^* A\,\vec{\boldsymbol{x}} \qquad \text{and} \qquad \langle Av, w \rangle = \langle v, A^* w \rangle.$$

For any matrix M we write $M^* = \overline{M^t}$, where M^t is the transpose of M, and $\overline{M^t}$ is the matrix obtained by complex conjugating every entry in M^t. (If M is a real matrix then $M^* = M^t$.)

Definition 15.1. Let A be an $n \times n$ complex matrix.

- A is *unitary* if $A^* A = I$.
- A is *normal* if $A^* A = AA^*$.
- A is *Hermitian* if $A^* = A$.
- A is *positive definite* if it is Hermitian and $\langle \vec{\boldsymbol{x}}, A\,\vec{\boldsymbol{x}} \rangle > 0$ for all nonzero vectors $\vec{\boldsymbol{x}}$.

The Spectral Theorem below guarantees that the definition for positive definiteness makes sense (since the theorem tells us that $\langle \vec{\boldsymbol{x}}, A\,\vec{\boldsymbol{x}} \rangle$ is a real number for Hermitian matrices A). We have the exact same definitions if A is real, but with slightly different names. For clarity let us repeat them.

Definition 15.2. Let A be an $n \times n$ real matrix.

- A is *orthogonal* if $A^t A = I$.
- A is *normal* if $A^t A = AA^t$.
- A is *symmetric* if $A^t = A$.
- A is *positive definite* if it is symmetric and $\langle \vec{\boldsymbol{x}}, A\,\vec{\boldsymbol{x}} \rangle > 0$ for all nonzero vectors $\vec{\boldsymbol{x}}$.

Lemma 15.3. *Every $n \times n$ matrix A can be written as $A = UTU^*$, where U is a unitary matrix and T is upper-triangular.*

Proof. Use the same proof as part (a) of Lemma 7.2 in Lecture 8. One can make sure U is unitary by demanding the eigenvectors we pick be of norm one. □

We now come to the Spectral Theorem, which together with the Jordan Canonical Form are the two main theorems in linear algebra. The Spectral Theorem applies only to normal matrices, and is purely an existence theorem (so it does not tell us how to compute anything). Consequently, or coincidentally, it has a much easier proof.

Theorem 15.4 (Spectral Theorem)**.** *Let A be an $n \times n$ complex matrix.*

(a) A is normal if and only if it is unitarily diagonalizable, i.e. $A = UDU^*$ for some unitary matrix U and diagonal matrix D.

(b) A is Hermitian if and only if it is unitarily diagonalizable with all eigenvalues real. Hence $A = UDU^*$ for some unitary matrix U and real diagonal matrix D.

(c) A is real symmetric if and only if it is unitarily diagonalizable with real eigenvalues and real eigenvectors. Hence $A = UDU^T$ for some orthogonal matrix U and real diagonal matrix D.

Proof. (a) The previous lemma tells us that $A = UTU^*$ for some unitary matrix U and upper-triangular matrix T. As A is normal, this tells us that $TT^* = T^*T$, forcing T to be diagonal.

(b) Let v be an eigenvector with eigenvalue λ. Then

$$\overline{\lambda}\langle v, v\rangle = \langle Av, v\rangle = \langle v, A^*v\rangle = \langle v, Av\rangle = \lambda\langle v, v\rangle.$$

so λ must be real.

(c) Part (b) tells us that every eigenvalue λ must be real. Since A is a real matrix, this implies that $\ker(A - \lambda I) > 0$, so it has positive real dimension. □

Corollary 15.5. *A Hermitian matrix has a basis of eigenvectors with real eigenvalues.*

Proof. Direct from the Spectral Theorem. □

The Spectral Theorem tells us that Hermitian matrices do not have generalized eigenvectors, so we do not have to think about Jordan Canonical Forms! In preparation for the next lecture we want to obtain a good basis of eigenvectors, called an orthonormal basis. Recall that the *norm* of a vector $\vec{v}$ is defined to be $\| \vec{v} \| := \langle \vec{v}, \vec{v}\rangle^{1/2}$.

Definition 15.6. Two vectors $\vec{v}$ and $\vec{w}$ are *orthogonal* if $\langle \vec{v}, \vec{w}\rangle = 0$. A set of vectors $v_1, \dots, v_n$ is *orthonormal* if it satisfies three conditions:

- $v_1, \dots, v_n$ forms a basis of F^n;
- each v_i has norm 1;
- the vectors are pairwise orthogonal.

Here is an important fact about eigenvectors. It tells us that eigenvectors for different eigenvalues are orthogonal.

Proposition 15.7. *Let A be a Hermitian matrix. For $j = 1, 2$, let v_j be an eigenvector for the eigenvalue λ_j, such that $\lambda_1 \neq \lambda_2$. Then v_1 and v_2 are orthogonal.*

Proof. As A is Hermitian $\langle v_1, Av_2\rangle = \langle Av_1, v_2\rangle$. Thus, as the eigenvalues of A are real,

$$\lambda_2\langle v_1, v_2\rangle = \lambda_1\langle v_1, v_2\rangle.$$

This implies $\langle v_1, v_2\rangle = 0$. □

How do we obtain an orthonormal basis from a given basis of eigenvectors?

Definition 15.8 (Gram-Schmidt Process)**.** Let $v_1, \dots, v_m$ be a set of linearly independent vectors in F^n. This iterative process produces an orthonormal set of vectors $b_1, \dots, b_m$ as follows.

- Set
$$u_1 = v_1 \qquad \text{and} \qquad b_1 = \frac{u_1}{\|u_1\|}.$$
- Now set
$$u_2 = v_2 - \frac{\langle v_2, u_1\rangle}{\langle u_1, u_1\rangle}u_1 \qquad \text{and} \qquad b_2 = \frac{u_2}{\|u_2\|}.$$
- Iterating, set
$$u_j = v_j - \frac{\langle v_j, u_1\rangle}{\langle u_1, u_1\rangle}u_1 - \cdots - \frac{\langle v_j, u_{j-1}\rangle}{\langle u_{j-1}, u_{j-1}\rangle}u_{j-1} \qquad \text{and} \qquad b_j = \frac{u_j}{\|u_j\|}.$$

It can be easily checked that $\langle b_i, b_j \rangle = 0$ for $i \neq j$.

Proposition 15.9. *A Hermitian matrix has an orthonormal basis of eigenvectors with real eigenvalues.*

Proof. We know that eigenvectors for different eigenvalues are already orthogonal, so we just need to apply the Grams-Schmidt Process to a basis for every eigenspace. ☐

Example 15.10. Consider the vectors

$$v_1 = (1, 2, 3, 0), \qquad v_2 = (1, 2, 0, 0), \qquad v_3 = (1, 0, 0, 1).$$

The Gram-Schmidt Process tells us that

$$u_1 = (1, 2, 3, 0), \qquad u_2 = (\frac{9}{14}, \frac{9}{7}, -\frac{15}{14}, 0), \qquad u_3 = (\frac{4}{5}, -\frac{2}{5}, 0, 1),$$

and

$$b_1 = \frac{1}{\sqrt{14}}(1, 2, 3, 0), \qquad b_2 = \frac{1}{3\sqrt{70}}(9, 18, -15, 0), \qquad b_3 = \frac{1}{3\sqrt{5}}(4, -2, 0, 5).$$

15.2 Positive Definite Matrices

The type of matrix that comes into play for mechanical oscillations are positive definite matrices. Henceforth let us note some facts for these kinds of matrices.

Proposition 15.11. *Let A be a positive definite $n \times n$ matrix.*

- *Every eigenvalue of A is positive.*
- *A is an invertible matrix.*
- *If C is another invertible $n \times n$ matrix, then C^*AC is also positive definite.*

Proof. (a) The Spectral Theorem tells us that any eigenvalue λ of A is real. Pick an eigenvector $\vec{v}$ for λ. Then $\langle \vec{v}, A\vec{v} \rangle > 0$. But

$$\langle \vec{v}, A\vec{v} \rangle = \langle \vec{v}, \lambda\vec{v} \rangle = \lambda \langle \vec{v}, \vec{v} \rangle$$

Since $\vec{v}$ is nonzero, this implies $\lambda > 0$.

(b) It suffices to show that $A = 0$. Suppose $\vec{x}$ satisfies $A\vec{x} = 0$. Then $\langle \vec{x}, A\vec{x} \rangle = 0$. By positive definiteness this forces $\vec{x} = 0$, as desired.

(c) Since C is invertible, any nonzero vector $\vec{x}$ satisfies $C\vec{x} \neq 0$. Therefore

$$\langle \vec{x}, C^*AC\vec{x} \rangle = \vec{x}^*C^*AC\vec{x} = (C\vec{x})^*AC\vec{x} = \langle C\vec{x}, AC\vec{x} \rangle > 0$$

by positive definiteness of A. ☐

Here is a criteria to check if a matrix is positive definite.

Proposition 15.12. *An $n \times n$ matrix A is positive definite if it is Hermitian and every eigenvalue is positive.*

Proof. Let $b_1, \ldots, b_n$ be an orthonormal basis of eigenvectors with positive eigenvalues $\lambda_1, \ldots, \lambda_n$ (the eigenvalues might not be distinct). Let $\vec{x}$ be a nonzero vector. Then it is easy to see that

$$\vec{x} = \langle \vec{x}, b_1 \rangle b_1 + \cdots + \langle \vec{x}, b_n \rangle b_n.$$

Thus a computation tells us

$$\langle \vec{\boldsymbol{x}}, A\,\vec{\boldsymbol{x}}\rangle = \lambda_1|\langle \vec{\boldsymbol{x}}, b_1\rangle|^2 + \cdots + \lambda_n|\langle \vec{\boldsymbol{x}}, b_n\rangle|^2.$$

Since the eigenvalues are positive, the expression above is nonnegative. It is zero exactly when $\langle \vec{\boldsymbol{x}}, b_j\rangle = 0$ for all j, implying $\vec{\boldsymbol{x}} = 0$. Therefore A is positive definite. □

Example 15.13. Consider the matrix

$$\begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}.$$

This matrix is symmetric with characteristic polynomial $(x-25)(x-9)$, so it is positive definite.

A key fact about positive definite matrices lies in the next theorem, of which we will not give a proof of, but we will use repeatedly in the next lecture.

Theorem 15.14. *Let A be a positive definite matrix. Then there is a unique positive definite square root $A^{1/2}$ of A. In fact, the square root can be constructed as follows.*

- *Let $b_1, \ldots, b_n$ be an orthonormal basis of eigenvectors with positive eigenvalues $\lambda_1, \ldots, \lambda_n$.*
- *Write*

$$U = \left[\; b_1 \;\middle|\; b_2 \;\middle|\; \cdots \;\middle|\; b_n \;\right] \qquad \text{and} \qquad D^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{\lambda_n} \end{bmatrix}.$$

- *Then $A^{1/2} = UD^{1/2}U^*$.*

Proof. Omitted; the existence follows from the Spectral Theorem, and the uniqueness is an elementary but slightly tricky argument that is not relevant to the course. □

Example 15.15. Again consider the positive matrix

$$A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

with eigenvalues 25 and 9, and corresponding orthonormal basis of eigenvectors

$$b_1 = \frac{1}{\sqrt{2}}(1,1) \qquad \text{and} \qquad b_2 = \frac{1}{\sqrt{2}}(1,-1).$$

Then

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad D^{1/2} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}, \qquad A^{1/2} = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}.$$

16 Lecture 16 – Application to Second Order ODEs Part 3

In this lecture we introduce the mechanical oscillation equation.

16.1 Oscillations

Definition 16.1. A *mechanical oscillation equation* is an equation

$$M\,\vec{\boldsymbol{x}}''(t) = -A\,\vec{\boldsymbol{x}}(t)$$

where M and A are positive definite matrices.

Proposition 16.2. *Any mechanical oscillation equation as above can be converted to*

$$\vec{\boldsymbol{y}}''(t) = -K\,\vec{\boldsymbol{y}}(t),$$

where $\vec{\boldsymbol{y}}(t) = M^{1/2}\,\vec{\boldsymbol{x}}(t)$ *and* $K = M^{-1/2}AM^{-1/2}$.

Proof. Computation. Note that everything written above makes sense by the previous lecture. □

Hence we now concentrate on understanding a mechanical oscillation equation of the form

$$\vec{\boldsymbol{y}}''(t) = -K\,\vec{\boldsymbol{y}}(t), \qquad \vec{\boldsymbol{y}}(0) = (a_1, \ldots, a_n), \qquad \vec{\boldsymbol{y}}'(0) = (b_1, \ldots, b_n).$$

(For convenience we will just let time start from zero.)

Let $b_1, \ldots, b_n$ be an orthonormal basis of eigenvectors of K with positive eigenvalues $\lambda_1, \ldots, \lambda_n$. By taking the dot product of the equation with b_j, one immediately sees that

$$\langle b_j, \vec{\boldsymbol{y}}(t)\rangle'' = -\lambda_j \langle b_j, \vec{\boldsymbol{y}}(t)\rangle.$$

Hence, by defining the *normal modes* $w_j = \langle b_j, \vec{\boldsymbol{y}}(t)\rangle$, one needs to solve

$$w_j''(t) = -\lambda_j w_j(t), \qquad w_j(0) = \langle b_j, \vec{\boldsymbol{y}}(0)\rangle, \qquad w_j'(0) = \langle b_j, \vec{\boldsymbol{y}}'(0)\rangle.$$

This is the Hooke's Law equation that we solved in Lecture 10! Recall that the solution to this equation is

$$w_j(t) = w_j(0)\cos(\sqrt{\lambda_j}t) + \frac{w_j'(0)}{\sqrt{\lambda_j}}\sin(\sqrt{\lambda_j}t).$$

The solution $\vec{\boldsymbol{y}}(t)$ can then be obtained by

$$\vec{\boldsymbol{y}}(t) = w_1(t)b_1 + \cdots + w_n(t)b_n,$$

and the actual solution we want is $\vec{\boldsymbol{x}}(t) = M^{-1/2}\,\vec{\boldsymbol{y}}(t)$.

Example 16.3. Consider the system of equation

$$\begin{aligned} 4x_1''(t) &= -8x_1(t) - 4x_2(t) \\ x_2''(t) &= -4x_1(t) - 5x_2(t) \end{aligned}$$

with initial conditions $x_1(0) = -1$, $x_2(0) = 1$, $x_1'(0) = 1/2$, $x_2'(0) = 2$. If we let

$$M = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \qquad A = \begin{bmatrix} 8 & 4 \\ 4 & 5 \end{bmatrix}, \qquad \vec{\boldsymbol{x}}(t) = (x_1(t), x_2(t)),$$

then the equation above is the mechanical oscillation equation

$$M\,\vec{\boldsymbol{x}}''(t) = -A\,\vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(0) = (-1,1), \qquad \vec{\boldsymbol{x}}'(0) = (\frac{1}{2},2).$$

We see that

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

so we do the conversion $\vec{\boldsymbol{y}}(t) = M^{1/2}\,\vec{\boldsymbol{x}}(t)$ and $K = M^{-1/2}AM^{-1/2}$ to get

$$K = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}, \qquad \vec{\boldsymbol{y}}(t) = (-2,1), \qquad \vec{\boldsymbol{y}}'(t) = (1,2).$$

Thus we now need to solve for

$$\vec{\boldsymbol{y}}''(t) = -K\,\vec{\boldsymbol{y}}(t), \qquad \vec{\boldsymbol{y}}(0) = (-2,1), \qquad \vec{\boldsymbol{y}}'(0) = (1,2).$$

The eigenvalues for K are 1 and 6, and we choose corresponding orthonormal basis of eigenvectors

$$b_1 = \frac{1}{\sqrt{5}}(-2,1), \qquad b_2 = \frac{1}{\sqrt{5}}(1,2).$$

By following the steps above one sees that

$$\vec{\boldsymbol{y}}(t) = \cos(t)(-2,1) + \frac{1}{\sqrt{6}}\sin(\sqrt{6}t)(1,2) = (-2\cos(t) + \frac{1}{\sqrt{6}}\sin(\sqrt{6}t), \cos(t) + 2\sin(\sqrt{6}t)).$$

The solution we seek is then $\vec{\boldsymbol{x}}(t) = M^{-1/2}\,\vec{\boldsymbol{y}}(t)$, so

$$\vec{\boldsymbol{x}}(t) = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sin(\sqrt{6}t) - 2\cos(t) \\ 2\sin(\sqrt{6}t) + \cos(t) \end{bmatrix} = (-\cos(t) + \frac{1}{2\sqrt{6}}\sin(\sqrt{6}t), \cos(t) + 2\sin(\sqrt{6}t)).$$

16.2 Driven Oscillations

Definition 16.4. A *driven mechanical oscillation equation* with *forcing term* $\vec{\boldsymbol{f}}(t)$ is an equation

$$M\,\vec{\boldsymbol{x}}''(t) = -A\,\vec{\boldsymbol{x}}(t) + \vec{\boldsymbol{f}}(t)$$

where M and A are positive definite matrices.

Proposition 16.5. *Any driven mechanical oscillation equation as above can be converted to*

$$\vec{\boldsymbol{y}}''(t) = -K\,\vec{\boldsymbol{y}}(t) + \vec{\boldsymbol{g}}(t),$$

where $\vec{\boldsymbol{y}}(t) = M^{1/2}\,\vec{\boldsymbol{x}}(t)$ *and* $K = M^{-1/2}AM^{-1/2}$ *and* $\vec{\boldsymbol{g}}(t) = M^{-1/2}\,\vec{\boldsymbol{f}}(t)$.

Proof. Computation. □

Just as before we now concentrate on understanding a mechanical oscillation equation of the form

$$\vec{\boldsymbol{y}}''(t) = -K\,\vec{\boldsymbol{y}}(t) + \vec{\boldsymbol{g}}(t), \qquad \vec{\boldsymbol{y}}(0) = (a_1,\ldots,a_n), \qquad \vec{\boldsymbol{y}}'(0) = (b_1,\ldots,b_n).$$

Let $b_1,\ldots,b_n$ be an orthonormal basis of eigenvectors of K with positive eigenvalues $\lambda_1,\ldots,\lambda_n$. By taking the dot product of the equation with b_j, one immediately sees that

$$\langle b_j, \vec{\boldsymbol{y}}(t)\rangle'' = -\lambda_j\langle b_j, \vec{\boldsymbol{y}}(t)\rangle.$$

Hence, by defining $w_j = \langle b_j, \vec{\boldsymbol{y}}(t)\rangle$ and $g_j(t) = \langle b_j, \vec{\boldsymbol{g}}(t)\rangle$, one needs to solve

$$w_j''(t) = -\lambda_j w_j(t) + g_j(t), \qquad w_j(0) = \langle b_j, \vec{\boldsymbol{y}}(0)\rangle, \qquad w_j'(0) = \langle b_j, \vec{\boldsymbol{y}}'(0)\rangle.$$

We have solved this equation using Duhamel's Formula in Lecture 14:

$$w_j(t) = w_j(0)\cos(\sqrt{\lambda_j}t) + \frac{w_j'(0)}{\sqrt{\lambda_j}}\sin(\sqrt{\lambda_j}t) + \frac{1}{\sqrt{\lambda_j}}\int_0^t \sin(\sqrt{\lambda_j}(t-s))g_j(s)\,ds.$$

The solution $\vec{\boldsymbol{y}}(t)$ can then be obtained by

$$\vec{\boldsymbol{y}}(t) = w_1(t)b_1 + \cdots + w_n(t)b_n,$$

and the actual solution we want is $\vec{\boldsymbol{x}}(t) = M^{-1/2}\,\vec{\boldsymbol{y}}(t)$.

There is a subtlety for driven mechanical oscillations: the normal modes $w_j(t)$ might demonstrate *resonance*. This happens when $g_j(t)$ models a periodic driving force in the form of some constant nonzero multiple of $\cos(\omega t + \phi_0)$. Thus, let us now look more closely at

$$\int_0^t \sin(\sqrt{\lambda_j}(t-s))\cos(\omega s + \phi_0)\,ds.$$

If we do integration using trigonometric identities, this equals

$$\frac{\cos(\omega t + \phi_0) - \cos(\sqrt{\lambda_j}t + \phi_0)}{2(\sqrt{\lambda_j} - \omega)} + \frac{\cos(\omega t + \phi_0) - \cos(-\sqrt{\lambda_j}t + \phi_0)}{2(\sqrt{\lambda_j} + \omega)}.$$

Notice that this function is defined everywhere except at $\omega = \sqrt{\lambda_j}$. If we perform the limit for the first term we see using L'Hopital's Rule that

$$\lim_{\omega \to \sqrt{\lambda_j}} \frac{\cos(\omega t + \phi_0) - \cos(\sqrt{\lambda_j}t + \phi_0)}{2(\sqrt{\lambda_j} - \omega)} = \frac{1}{2}t\sin(\sqrt{\lambda_j}t + \phi_0).$$

Thus $w_j(t)$ grows without bound if $\omega = \sqrt{\lambda_j}$. We now make a definition.

Definition 16.6. If $g_j(t)$ is a nonzero constant multiple of $\cos(\omega t + \phi_0)$, then the driven mechanical oscillation equation demonstrates *resonance* at $\omega = \sqrt{\lambda_j}$.

Resonance is a terrifying thing. For example, it caused the collapse of the Tacoma Narrows Bridge. A bad experiment to try at home is to get an oscillator to produce sound at the natural frequency of a wine glass.

Example 16.7. Consider

$$\vec{\boldsymbol{y}}''(t) = -K\,\vec{\boldsymbol{y}}(t) + \vec{\boldsymbol{f}}(t), \qquad \vec{\boldsymbol{y}}(0) = (0,0), \qquad \vec{\boldsymbol{y}}'(0) = (1,2).$$

with

$$K = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}, \qquad \vec{\boldsymbol{f}}(t) = \cos(\omega t)(-2,1).$$

As before the eigenvalues for K are 1 and 6. After some computations we will get the ODEs for the normal modes:

$$w_1''(t) = -w_1(t) + \sqrt{5}\cos(\omega t), \qquad w_1(0) = w_1'(0) = 0,$$
$$w_2''(t) = -6w_2(t), \qquad w_2(0) = 0, \qquad w_2'(0) = \sqrt{5}.$$

These ODE's can be solved using the equations above to get $\vec{\boldsymbol{x}}(t)$. For this ODE one sees that there is resonance at $\omega = 1$, but not at $\omega = \sqrt{6}$.

17 Lecture 17 – The Euler-Lagrange Equation

This lecture discusses the Euler-Lagrange Equation, which gives stationary solutions for special kinds of functionals important in physical applications. We will assume all functions are nice enough, and focus on examples. In particular, we will not sketch a proof of the Euler-Lagrange Equation (it is a long but uncomplicated exercise in real analysis).

17.1 Statement of the Euler-Lagrange Equation

We start by fixing notations. Let $\vec{\boldsymbol{x}}(t)$ be any t-valued vector, and let $\vec{\boldsymbol{v}}(t) = \vec{\boldsymbol{x}}'(t)$. Recall that, if $\vec{\boldsymbol{x}}(t) = (x_1(t), \ldots, x_n(t))$, then

$$\nabla_{\vec{\boldsymbol{x}}} f\left(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{v}}(t), t\right) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right),$$

and similarly for $\nabla_{\vec{\boldsymbol{v}}} f\left(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{v}}(t), t\right)$.

Theorem 17.1 (Euler-Lagrange)**.** *If a function $\vec{\boldsymbol{x}}(t)$ minimizes or maximizes a functional*

$$I(\vec{\boldsymbol{x}}) = \int_a^b f\left(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{v}}(t), t\right)\, dt$$

then

$$\nabla_{\vec{\boldsymbol{x}}} f\left(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{v}}(t), t\right) = \frac{d}{dt} \nabla_{\vec{\boldsymbol{v}}} f\left(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{v}}(t), t\right). \qquad \square$$

Corollary 17.2. *If f depends only on $\vec{\boldsymbol{x}}(t)$ and $\vec{\boldsymbol{v}}(t)$, and not t, then the Euler-Lagrange Equation reduces to*

$$\frac{d}{dt}\left[\vec{\boldsymbol{x}}'(t) \nabla_{\vec{\boldsymbol{v}}} f(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{x}}'(t)) - f(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{x}}'(t))\right] = 0.$$

Proof. Observe that

$$\frac{d}{dt}\left[\vec{\boldsymbol{x}}'(t) \nabla_{\vec{\boldsymbol{v}}} f(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{x}}'(t)) - f(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{x}}'(t))\right] = \vec{\boldsymbol{x}}'(t)\left[\frac{d}{dt} \nabla_{\vec{\boldsymbol{v}}} f\left(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{v}}(t), t\right) - \nabla_{\vec{\boldsymbol{x}}} f\left(\vec{\boldsymbol{x}}(t), \vec{\boldsymbol{v}}(t), t\right)\right].$$

The corollary now follows immediately from the Euler-Lagrange Equation. $\square$

17.2 Examples

Example 17.3. Let us try to find $f : [0,1] \longrightarrow \mathbb{R}$ such that $f(0) = 0$ and $f(1) = 1$ and such that the arc length

$$\int_0^1 \sqrt{1 + (f'(t))^2}\, dx$$

is minimized. Applying the Euler-Lagrange Equation tells us that

$$0 = \frac{d}{dt}\left(\frac{f'(t)}{\sqrt{1 + (f'(t))^2}}\right)$$

so

$$\frac{f'(t)}{\sqrt{1 + (f'(t))^2}} = c$$

for some constant c. An elementary manipulation tells us that

$$f'(t) = \frac{c}{\sqrt{1-c^2}} := a,$$

so $f(t) = at + b$ for some other constant b. Since $f(0) = 0$ and $f(1) = 1$, this means $f(t) = t$ is the minimizing function.

Example 17.4. Consider the functional

$$I(\vec{\boldsymbol{x}}) = \frac{1}{2}\int_0^1 \| \vec{\boldsymbol{x}}'(t)\|^2 \, dt.$$

The Euler-Lagrange Equation in this case is

$$0 = \frac{d}{dt}\,\vec{\boldsymbol{x}}'(t),$$

or $\vec{\boldsymbol{x}}''(t) = 0$. If we fix boundary conditions $\vec{\boldsymbol{x}}(0) = \vec{\boldsymbol{\alpha}_0}$ and $\vec{\boldsymbol{x}}(1) = \vec{\boldsymbol{\alpha}_1}$, this means the solution to the Euler-Lagrange Equation is the straight line through $\vec{\boldsymbol{\alpha}_0}$ and $\vec{\boldsymbol{\alpha}_1}$.

Example 17.5. Consider the functional

$$I(x) = \int_0^1 t\left(x'(t)\right)^2 \, dt.$$

The Euler-Lagrange Equation is

$$0 = 2tx'(t),$$

so we require $tx'(t) = c$. Hence

$$x(t) = c\ln t + b.$$

Let us now impose boundary conditions $f(0) = \alpha$ and $f(1) = \beta$. For $x(t)$ to be defined at 0 one must then have $c = 0$, but this gives us no possible values for b. Hence there is no minimizing function for this Euler-Lagrange Equation.

Example 17.6. Consider the functional

$$I(x) = \int_0^L (x'(t))^2 - (x(t))^2 - x(t)\sin t \, dt$$

with boundary conditions $x(0) = x(\pi/4) = 0$. The Euler-Lagrange Equation is

$$-2x(t) - \sin t = \frac{d}{dt}(2x'(t))$$

or in other words $2x'' + 2x = -\sin t$. By Duhamel's Formula, or by observation,

$$x(t) = c_1 \sin t + c_2 \cos t + \frac{1}{4} t\cos t.$$

Applying the boundary conditions tells us that

$$x(t) = \frac{1}{4} t\cos t - \frac{\pi}{16}\sin t.$$

17.3 The Brachistochrone Problem

Let $p_0 = (x_0, y_0) = (0, 0)$, and pick a point $p_1 = (x_1, y_1)$ in the bottom right quadrant of the plane. The Brachistochrone Problems asks the following.

Question. Find the shape of the wire $y(x)$ that minimizes the time a bead slides from p_0 to p_1.

Note that, by physical reasons,

- $y(x)$ and $y'(x)$ should be nonpositive, and
- the only force acting on the bead is gravity g.

The first step is to derive the functional that models travel time along $y(x)$. Let m be the mass of the bead. When the bead slides to height $y < 0$, the potential energy that has been converted to kinetic energy is

$$-mgy.$$

As the formula for kinetic energy is $mv^2/2$, this implies the speed of the bead at this height is

$$v = \sqrt{-2gy},$$

Recall that the time to travel from p_0 to p_1 is given by the path integral

$$\int_{p_0}^{p_1} \frac{1}{v(x)}\, ds,$$

where s is the arc length. Since

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + y'(x)^2}\, dx$$

tells us that the time it takes to move along this interval is

$$\int_0^{x_1} \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{-2gy(x)}}\, dx.$$

This is the functional that we are looking to minimize. By Corollary 17.2,

$$y'(x)\frac{y'(x)}{\sqrt{-2gy(x)}\sqrt{1 + (y'(x))^2}} - \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{-2gy(x)}} = c$$

for some constant c. A manipulation gives us the equation

$$y(x)[1 + (y'(x))^2] = C$$

for some other constant C. Note that C is necessarily a negative constant. Recalling that y' should be a decreasing function, this implies

$$y'(x) = -\left(\frac{C - y}{y}\right)^{1/2}.$$

By separation of variables

$$dx = -\left(\frac{y}{C - y}\right)^{1/2} dy.$$

We now change variables $y = C \sin^2 \phi$, implying

$$dx = -2C \sin^2 \phi\, d\phi = -C(1 - \cos 2\phi)\, d\phi.$$

After integrating

$$x = -\frac{C}{2}(2\phi - \sin 2\phi) + B.$$

By substituting $x_0 = 0$, this implies $B = 0$. Hence, if we let $r = -C/2$ and $\theta = 2\phi$, then a parametric equation for $y(x)$ is

$$x(\theta) = r(\theta - \sin\theta), \qquad y(\theta) = -r(1 - \cos\theta).$$

This is the equation for a cycloid! In other words, the cycloid minimizes time (and not the straight line or some sort of parabola/hyperbola, disagreeing with common intuition).

18 Lecture 18 – Final Review Part 1

18.1 Final Review Part 1

A first review of lectures 1 to 17 was given during this lecture. You should be comfortable with the topics below; this list is also repeated in the next lecture.

Matrices

- Compute inverses using cofactor expansion or the Cayley-Hamilton Theorem
- Compute powers and exponentials using Jordan Canonical Form or Buchheim's Algorithm
- Exponentiation formulas for nilpotent, diagonalizable, and small matrices
- Positive definite matrices

First Order Constant Linear ODEs

- An application of invariant subspaces
- Reduction of order
- Linearization and stability

Higher Order Linear ODEs

- Tricks for finding linearly independent solutions
- Duhamel's Formula
- Oscillations and resonance

Other Aspects of ODEs

- The Euler-Lagrange Equation

19 Lecture 19 – Final Review Part 2

19.1 Final Review Part 2

A second review of lectures 1 to 17 was given during this lecture. You should be comfortable with the topics below; this list is also repeated in the next lecture.

Matrices

- Compute inverses using cofactor expansion or the Cayley-Hamilton Theorem
- Compute powers and exponentials using Jordan Canonical Form or Buchheim's Algorithm
- Exponentiation formulas for nilpotent, diagonalizable, and small matrices
- Positive definite matrices

First Order Constant Linear ODEs

- An application of invariant subspaces
- Reduction of order
- Linearization and stability

Higher Order Linear ODEs

- Tricks for finding linearly independent solutions
- Duhamel's Formula
- Oscillations and resonance

Other Aspects of ODEs

- The Euler-Lagrange Equation

20 Lecture 20 – Final

20.1 Final

The final was given during this lecture; see subsection 21.10. Below is a non-exhaustive list of things that you should know for the final.

Matrices

- Compute inverses using cofactor expansion or the Cayley-Hamilton Theorem
- Compute powers and exponentials using Jordan Canonical Form or Buchheim's Algorithm
- Exponentiation formulas for nilpotent, diagonalizable, and small matrices
- Positive definite matrices

First Order Constant Linear ODEs

- An application of invariant subspaces
- Reduction of order
- Linearization and stability

Higher Order Linear ODEs

- Tricks for finding linearly independent solutions
- Duhamel's Formula
- Oscillations and resonance

Other Aspects of ODEs

- The Euler-Lagrange Equation

21 Syllabus, Homework, and Exams

21.1 Syllabus

Course: Math 240 (Calculus, Part III)
Instructor: Yao-Rui
Email: yeya@sas.upenn.edu
Webpage: Penn Canvas
Lectures: MTWR, 13:00 – 15:10, 3W2 DRL, from 2019-07-05 to 2019-08-09
First Day of Class: 2019-07-08
Office Hours: Immediately after each lecture

Overview. This course covers applications of linear algebra to solving differential equations. We will focus our study to first and second order ordinary differential equations (ODEs). The goal by the end of this course is not only to know how to solve various kinds of differential equations, but also to gain a working knowledge of linear algebra.

Prerequisites. I will assume Calculus I and II. This varies from institution to institution, but here at Penn this means the basics of univariate and multivariate calculus. One definitely has to be comfortable with elementary computations in calculus, but we will not use some stuff usually covered at the end of Calculus II (Stokes' Theorem comes to mind).

Text. I will not be following any textbook closely. Here are some references that I will be using to help develop the course.

- Past Calculus Finals, from the University of Pennsylvania.
- Differential Equations and Linear Algebra, by Stephen W. Goode and Scott A. Annin.
- Ordinary Differential Equations, by Wolfgang Walter.
- Linear Algebra, by Kenneth Hoffman and Ray Kunze.

Please do not spend money buying math books that are unreasonably expensive. I will post everything we need for this class on the course webpage.

Grading. The course will be graded accordingly:

- Five Homework (25 points),
- One Midterm (35 points),
- One Final (40 points),

for a maximum of 100 points. In addition, there is an opportunity to submit an extra credit assignment. Due dates are summarized below.

- Homework 1: July 11
- Homework 2: July 17
- Homework 3: July 23
- Midterm: July 25
- Homework 4: July 31
- Homework 5: August 6
- Final: August 8
- Extra Credit: Any lecture of your choice

Homework must be turned in, hardcopy, at the start of the class, and late homework will not be accepted. There will not be any makeup exams; the final will count for 75 points if you have to miss the midterm, and I hope no one misses the final. The course grade will be curved appropriately.

More about the Homework. This course emphasizes computational techniques heavily, so one should practice as much as one needs in order to be comfortable with the material. Thus the homework constitutes a significant part of the grade. The Math 240 Past Finals from the University of Pennsylvania, from Fall 2014 onwards, is another good place to get more practice problems.

More about the Exams. Both exams will be written, and you will have the entire lecture to work on it. As I do not believe in memorizing equations, you are allowed one single-sided letter-sized cheat sheet for each exam. The exams will be similar in spirit to the homework.

More about the Extra Credit. If you choose to do the extra credit assignment, you can turn it in during any lecture before (but *not* including) the final one. I may ask you to present your solutions during office hours, so be prepared.

Collaboration and Academic Integrity. If you have any questions or difficulties with anything at any point during the course, please discuss them with your classmates or talk to me during office hours. Working together on the homework is encouraged, but anything that is handed in must be written up individually. Any violation of academic integrity will result in serious consequences.

Tentative Schedule. Here is a planned week-by-week schedule of lectures.

- Week 1: Introduction; Matrices and Linear Maps; Eigenstuff.
- Week 2: Matrix Exponentials; First Order ODEs.
- Week 3: First Order ODEs Continued; Equilibrium Points; Review; Midterm.
- Week 4: Flows; Second Order ODEs.
- Week 5: The Euler-Lagrange Equation; Review; Final.

21.2 Homework 1

This homework reviews lectures 1 and 2. The last problem is optional and not for credit.

Problem 1. Solve the radioactive decay equation $N' = -kN$ with $N(0) = 10$. What is the value of k if $N(3) = 500$?

Problem 2. Solve $x^2x' = 5x^3 + e^{-t}$ with initial condition $x(0) = 2$.

Problem 3. Find the general solution of $x' = -x + \sin t$. Also determine an initial condition $x(\pi/2) = \iota$ such that the resulting solution is periodic.

Problem 4. Find the general solution of $x' = x^2 - x - 2$.

Problem 5. This problem is related to row echelon forms.

(a) Use row reduction to solve the system of equations

$$\begin{aligned} x + 2y + 3z &= 6 \\ 2y + 3z &= 5 \\ x + 4y + 6z &= 11. \end{aligned}$$

(b) Do the same with

$$\begin{aligned} x + 2y + 3z &= 6 \\ 2y + 3z &= 5 \\ x + 4y + 7z &= 11. \end{aligned}$$

(c) Do the same with

$$\begin{aligned} x + 2y + 3z &= 6 \\ 2y + 3z &= 5 \\ x + 4y + 6z &= 10. \end{aligned}$$

Problem 6. Find the rank of the matrix

$$\begin{bmatrix} 3 & -1 & 1 \\ -7 & 4 & t \\ 2 & 1 & 4 \end{bmatrix}.$$

Your answer should depend on t.

Problem 7. Compute the inverse of the matrix

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

Problem 8. Use Cramer's Rule to solve the system of equations

$$\begin{aligned} 2x + 3y - z &= 1 \\ 4x + y - 3z &= 8 \\ 3x - 2y + 5z &= 21. \end{aligned}$$

Problem *. The *Hilbert matrix* is the $n \times n$ matrix H_n with $(i + j - 1)^{-1}$ in the (i, j)-entry. For example,

$$H_4 = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{bmatrix}.$$

In this problem we will compute $\det H_n$, and show that it is the reciprocal of an integer. (In fact one can also compute the inverse of the Hilbert matrix and show that it has integer entries.)

(a) Define $c_n = \prod_{j=1}^{n} j!$. Show that

$$\det H_n = \frac{c_n^4}{c_{2n-1}}$$

by doing the following steps.

- Subtract the last row of H_n from every row above it.
- Factor out all common terms from each row and column.
- Subtract the last column of H_n from every column before it.
- Factor out all common terms from each row and column.
- Do induction on n.

(b) Show that

$$\frac{c_{2n-1}}{c_n^4} = n! \prod_{j=1}^{2n-1} \binom{j}{\lfloor j/2 \rfloor}$$

by induction on n. Hence $\det H_n$ is the reciprocal of an integer.

21.3 Homework 2

This homework reviews lectures 3 to 5. The last problem is optional and not for credit.

Problem 1. Compute the rank and nullity of the matrix

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

Also find the image and kernel of this matrix. (Assume we are working with real numbers.)

Problem 2. Here are two problems on the interplay between linear maps and matrices.

(a) Write down the linear map with standard matrix

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

Also write down five distinct invariant subspaces of this linear map, with explanation.

(b) Compute the change of basis matrix from $(3, 1), (-2, 1)$ to $(2, 1), (1, 4)$.

Problem 3. Consider the following subsets. Explain whether they are subspaces or not.

(a) $S = \{(x, y) \text{ in } \mathbb{R}^2 \text{ satisfying } y = x^2\}$.

(b) $S = \{(a + 3b + c, b, 0) \text{ with } a, b, c \text{ real numbers}\}$.

(c) $S = \{(x, y, z, w) \text{ in } \mathbb{R}^4 \text{ satisfying } x \le y \le z\}$.

Problem 4. Let $\mathcal{P}_n$ be the $(n + 1)$-dimensional vector space of polynomials with real coefficients having degree at most n. Consider the linear map $\varphi \colon \mathcal{P}_2 \longrightarrow \mathcal{P}_3$ such that

$$\varphi(1 + x) = x^3 - 2x, \qquad \varphi(2 - x^2) = 2x^3 + x^2 - 2, \qquad \varphi(x^2 + x) = -x^2 - 2x + 1.$$

Compute the matrix of φ with respect to the bases $1, x, x^2$ and $1, x, x^2, x^3$.

Problem 5. Find the eigenvalues and eigenspaces for the matrix

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}.$$

For each eigenspace, write down a basis of eigenvectors.

Problem 6. Let a and b be complex numbers. Compute the eigenvalues and eigenspaces of

$$\begin{bmatrix} a & b \\ \bar{b} & a + b + \bar{b} \end{bmatrix}.$$

For example, one can check that

$$\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix},$$

with $a = 1$ and $b = i$, has eigenvalues 2 and 0, with associated eigenspaces the span of $(i, 1)$ and $(-i, 1)$.

Problem 7. Here are two conceptual problems.

(a) Let A be an 8×8 matrix with $\text{rank}(A) = 5$. Show that $\text{rank}(A^2) \geq 2$.

(b) Show that any $n \times n$ matrix, with n odd, must have a real eigenvalue.

Problem 8. Show that the matrix

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

has exactly one eigenvalue. Hence or otherwise compute e^{2A}.

Problem *. In this problem we outline how one can model rotations using matrices.

(a) Using the basis vectors $(1,0)$ and $(0,1)$ in $\mathbb{R}^2$, write down a 2×2 matrix R_θ that models a θ-degree rotation counterclockwise about the origin. Show that $R_{\theta_1} R_{\theta_2} = R_{\theta_2} R_{\theta_1}$. What does this equality mean geometrically?

(b) Using the basis vectors $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ in $\mathbb{R}^3$, write down three 3×3 matrices that models a θ-degree rotation counterclockwise about the x, y, and z axes respectively. If M and N are 3-dimensional rotation matrices, is it still true that $MN = NM$?

(c) Show that, in both $\mathbb{R}^2$ and $\mathbb{R}^3$, the composition of two rotations that fixes the origin is still a rotation.

21.4 Homework 3

This homework reviews lectures 6 to 8. The last problem is optional and not for credit.

Problem 1. Compute the Jordan Canonical Forms for

$$\begin{bmatrix} -3 & -2 & 1 \\ 0 & -4 & 0 \\ 1 & -2 & -3 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 4 \\ 0 & 0 & -2 \end{bmatrix}.$$

Problem 2. Compute the Jordan Canonical Form for

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

Use this to compute e^A and A^{2019}. Your answers to e^A and A^{2019} should not be imaginary.

Problem 3. Suppose A is a 3 by 3 matrix with eigenvalues $1, i, -i$.

(a) Show that $A^{80085} = A$.

(b) Write down two such matrices A with real entries.

Problem 4. The Fibonacci Numbers is the sequence $1, 1, 2, 3, 5, 8, 13, \ldots$ defined by the recurrence

$$F_n = F_{n-1} + F_{n-2}, \qquad F_1 = F_2 = 1.$$

Note that we can write the recurrence in matrix notation as

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

Let A be the 2×2 matrix above.

(a) Show that

$$\begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

(b) Compute the Jordan Canonical Form for A, and use this to write down a formula for A^k.
(c) Use the above two parts to write down a formula for F_n. The formula should involve the golden ratio.

Problem 5. Compute A^{1438} for

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

(Although the JCF for this matrix was written down in lecture, you still need to show work.)

Problem 6. Find the solution to the equation

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(0) = \vec{\boldsymbol{x}}_0,$$

where

$$A = \begin{bmatrix} -5 & 3 & 2 \\ -8 & 5 & 4 \\ -4 & 3 & 3 \end{bmatrix}.$$

Problem 7. Find the solution to the equation

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t) + (1,1,1), \qquad \vec{\boldsymbol{x}}(1) = (2,5,3),$$

where

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

Problem 8. Solve the differential equation

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(0) = (1,1,1,0,0,0,0,0),$$

where A is the 8×8 matrix given by

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

Problem *. A conic section is the solutions of an equation of the form

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

on the Cartesian plane. One can write this equation in terms of a matrix as

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} D & E \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + F = 0.$$

In this problem we seek to classify all conic sections.

(a) First suppose $B = 0$ for simplicity. Give a criteria, depending on A and C, to determine if the conic section is a parabola, hyperbola, or ellipse.

(b) Now suppose $B \neq 0$. Show that

$$\begin{bmatrix} A & B/2 \\ B/2 & C \end{bmatrix}$$

is diagonalizable, and give a criteria depending on the determinant of this matrix to determine if the conic section is a parabola, hyperbola, or ellipse.

(c) Show that if the conic section is a circle, then $A = C$ and $B = 0$.

(d) Determine what kind of conic section

$$x^2 + 2y^2 + 2xy + x - 5y = 0$$

is, and sketch its graph.

21.5 Midterm Review Problems

This problem set reviews key ideas relevant to the midterm. The last problem is optional.

Problem 1. Find the general solution to

$$t^3 x' + t^2 x - x^2 = 2t^4.$$

Problem 2. Consider the following matrix.

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

(a) Write down the linear map of this matrix, assuming it is a standard matrix.

(b) Find the rank and nullity of this matrix.

(c) Compute the image and kernel of this matrix.

(d) Is this matrix invertible?

Problem 3. Consider the matrix

$$\begin{bmatrix} 2 & k & 0 \\ 0 & 3 & 0 \\ 0 & 0 & k \end{bmatrix}.$$

(a) Write down the linear map of this matrix, assuming it is a standard matrix.

(b) Find the rank and nullity of this matrix. Your answer should depend on k.

(c) For which values of k does the inverse matrix exist?

(d) For which values of k is the matrix diagonalizable?

Problem 4. Consider the matrix

$$A = \begin{bmatrix} 5 & -1 \\ 4 & 1 \end{bmatrix}.$$

Find the solution to

$$\vec{x}'(t) = A\,\vec{x}(t) + (2, t), \qquad \vec{x}(0) = (1, 1).$$

Problem 5. Compute e^{tA} for

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 4 & -1 \\ 0 & 1 & 2 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -2 & 4 & 2 & 0 \\ -1 & 0 & 4 & 1 \\ -2 & 0 & 0 & 4 \end{bmatrix}.$$

Problem 6. Compute e^{A} for

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

Problem 7. Consider the system of differential equations

$$\frac{dx}{dt} = (x^2 + y)(1 - x),$$
$$\frac{dy}{dt} = 3x - y.$$

Find the equilibrium points, and discuss the kind of stability around the equilibrium points. Do the same with the system

$$\frac{dx}{dt} = (2 + x)(y - x),$$
$$\frac{dy}{dt} = y(2 + x - x^2).$$

Problem 8. The *minimal polynomial* of an $n \times n$ matrix A is the smallest degree polynomial of the form

$$m(x) = x^d + c_{d-1}x^{d-1} + \cdots + c_1 x + c_0$$

such that $m(A) = 0$. This problem tells us that the minimal polynomial is closely related to the Jordan canonical form of A.

(a) Show that

$$m(x) = \prod_{\substack{\text{eigenvalues} \\ \lambda \text{ of } A}} (x - \lambda I)^{j_\lambda}$$

where j_λ is the maximum size among all the Jordan blocks associated to λ.

(b) Explain why the minimal polynomial of A divides the characteristic polynomial of A.

Problem *. Let $\boldsymbol{b}$ and $\boldsymbol{v}_0$ be two nonzero vectors in $\mathbb{R}^3$. In this problem we will solve the rotation equation

$$\vec{\boldsymbol{x}}'(t) = \boldsymbol{b} \times \vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(0) = \boldsymbol{v}_0,$$

where the multiplication symbol above denotes the cross product.

(a) Rewrite the rotation equation in the form $\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t)$ for some square matrix A.

(b) Show that the eigenvalues of A are 0, $i\|\boldsymbol{b}\|$, $-i\|\boldsymbol{b}\|$.

(c) Show that

$$e^{tA} = \frac{1}{\|\boldsymbol{b}\|^2}A^2 - \frac{\cos(\|\boldsymbol{b}\|t)}{\|\boldsymbol{b}\|^2}A^2 + \frac{\sin(\|\boldsymbol{b}\|t)}{\|\boldsymbol{b}\|}A + I.$$

Sylvester's Formula will be useful here.

(d) Show that the solution to the rotation equation is

$$\vec{x}(t) = \frac{1}{\|\boldsymbol{b}\|^2}(\boldsymbol{b} \times (\boldsymbol{b} \times \boldsymbol{v}_0)) - \frac{\cos\|\boldsymbol{b}\|}{\|\boldsymbol{b}\|^2}(\boldsymbol{b} \times (\boldsymbol{b} \times \boldsymbol{v}_0)) + \frac{\sin\|\boldsymbol{b}\|}{\|\boldsymbol{b}\|}(\boldsymbol{b} \times \boldsymbol{v}_0) + \boldsymbol{v}_0.$$

(e) Let

$$\boldsymbol{v}_0^\perp = \boldsymbol{v}_0 - \frac{\boldsymbol{v}_0 \cdot \boldsymbol{b}}{\|\boldsymbol{b}\|^2}\boldsymbol{b}$$

be the orthonormal projection of $\boldsymbol{v}_0$ onto $\boldsymbol{b}$. Consider the set of orthonormal basis

$$u_1 = \frac{1}{\|\boldsymbol{b}\|}\boldsymbol{b}, \qquad u_2 = \frac{1}{\|\boldsymbol{v}_0^\perp\|}\boldsymbol{v}_0^\perp, \qquad u_3 = \frac{1}{\|u_1 \times u_2\|}(u_1 \times u_2).$$

Show that the solution to the rotation equation can be rewritten as

$$\vec{x}(t) = \frac{\boldsymbol{b} \cdot \boldsymbol{v}_0}{\|\boldsymbol{b}\|}u_1 + \cos(\|\boldsymbol{b}\|t)\|\boldsymbol{v}_0^\perp\|\, u_2 + \sin(\|\boldsymbol{b}\|t)\|\boldsymbol{v}_0^\perp\|\, u_3.$$

For this part you may find the identity $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}$ useful.

(f) Finally, draw a quantitative graph of the solution $\vec{x}(t)$. It should be a circle about the axis through $\boldsymbol{b}$, with some specified radius and distance from the origin.

21.6 Midterm

The midterm consists of eight problems and an extra credit problem.

Problem 1. Find the general solution to

$$x'(t) = (x(t) - t)^2 + 1.$$

Your answer should have a constant C somewhere that is determined by the initial condition.

Problem 2. Let k be a real number.

(a) Find all k such that there exists a 3×3 real matrix A having:
- eigenvalue 1 with eigenvectors $(1, 1, 0)$ and $(1, 0, 1)$, and
- eigenvalue 3 with eigenvector $(2, k, 3)$.

(b) If such an A in part (a) exists, write down an example of it. Entries should involve k.

Problem 3. I will advise you not to compute the characteristic polynomial for this problem.

(a) Show that an $n \times n$ square matrix A has 0 has an eigenvalue if $\operatorname{rank}(A) < n$.

(b) Find all real numbers k such that the following matrix has 0 as an eigenvalue.

$$\begin{bmatrix} 0 & 1 & 0 & 1 & k^2 \\ 1 & 0 & 12345 & 0 & 3k \\ 0 & 0 & \ln 6789 & 0 & 0 \\ 2 & 1 & 2 & 2 & -6 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

Problem 4. Consider the matrix

$$A = \begin{bmatrix} 1 & 0 & k \\ 0 & 4 & 0 \\ 0 & 0 & k \end{bmatrix}.$$

(a) For which values of the real number k is A diagonalizable?

(b) For the values of k where A is not diagonalizable, compute A^{2019}. Your answer should not involve k.

Problem 5. Let $\theta = \pi/10$ throughout this problem. Consider the matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

(a) Show that

$$A^n = \begin{bmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{bmatrix}$$

for any positive integer n.

(b) Find the solution to

$$\vec{\boldsymbol{x}}'(t) = A^5\,\vec{\boldsymbol{x}}(t), \qquad \vec{\boldsymbol{x}}(0) = (2,3).$$

Your answer should not involve θ nor imaginary numbers.

Problem 6. Compute e^A for

$$A = \begin{bmatrix} 4 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix}.$$

Problem 7. Consider the matrix

$$A = \begin{bmatrix} 3 & -18 \\ 2 & -9 \end{bmatrix}.$$

Find the solution to

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t) + (1,t), \qquad \vec{\boldsymbol{x}}(0) = (1,1).$$

Problem 8. Consider the system of differential equations

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= x^4 - 5x^2 + 4. \end{aligned}$$

Find the equilibrium points, and discuss the kind of stability around the equilibrium points.

Extra Credit. Let A and B be two $n \times n$ matrices with $AB = BA$.

(a) Show that there exists an invertible matrix P such that PAP^{-1} and PBP^{-1} are both upper triangular. [Hint: Show that A has an eigenvalue λ, and argue that you can somehow consider B on the subspace $\ker(A - \lambda I)$.]

(b) If A is diagonalizable, show that there exists an invertible matrix P such that PAP^{-1} and PBP^{-1} are both diagonal.

21.7 Homework 4

This homework reviews lectures 9, 10, and 13. The last problem is optional and not for credit.

Problem 1. Consider the system of differential equations

$$\begin{aligned} \frac{dx}{dt} &= -y^3 + x, \\ \frac{dy}{dt} &= x^3 - y. \end{aligned}$$

(a) Find the equilibrium points.
(b) Linearize the system of differential equations around each equilibrium point.
(c) Discuss the kind of stability around the equilibrium points.

Problem 2. Consider

$$\frac{dx}{dt} = -x^3 e^{xy} - 2y,$$
$$\frac{dy}{dt} = 2xe^{xy} + x^2 y e^{xy}.$$

Show that this system of ODE is Hamiltonian, and compute the general solution curve.

Problem 3. Consider

$$\frac{dx}{dt} = -x\sin(y) + 2y,$$
$$\frac{dy}{dt} = -\cos(y).$$

Show that this system of ODE is Hamiltonian, and compute the general solution curve.

Problem 4. Find the general solution to $y^{(6)} - 2y^{(5)} + 5y^{(4)} = 0$, where $y^{(n)}$ means taking the n^{th}-derivative of y.

Problem 5. Find a particular solution to $y''(t) - 4y(t) - 12y(t) = 2t^3 - t + 3$ by guessing a solution of the form $y(t) = at^3 + bt^2 + ct + d$. You do not have to find the general solution.

Problem 6. Find the general solution to $x^3 y''' + xy' - y = 0$.

Problem 7. Solve for the general solution of

$$\frac{dy}{dx} = -3\frac{y}{x} - y^{\frac{3}{2}} x^{\frac{1}{2}}, \qquad x > 0$$

by following the steps below.

(a) Consider the change of variables $u = -\ln x$ and $v = x^3 y(x)$. Show that

$$v = e^{-3u} y(e^{-u}),$$

and, after letting $z = e^{-u}$, use the chain rule to deduce that

$$\frac{dv}{du} = -3z^3 y - z^4 \frac{dy}{dz}.$$

(b) By observing that z is in fact just x, show that

$$\frac{dv}{du} = v^{3/2}.$$

(c) Use the above ODE to find the general formula for $y(x)$.

Problem 8. Find the solution curves of

$$\frac{dy}{dx} = 4x^2 y^2 + x^5 y^3$$

by following the steps below.

(a) Consider the change of variables $u = \ln x$ and $v = x^3 y(x)$. Use the chain rule to deduce that

$$\frac{dv}{du} = 3x^3 y + x^4 \frac{dy}{dx}.$$

(b) Show that

$$\frac{dv}{du} = 3v + 4v^2 + v^3.$$

Then use partial fraction decomposition, or otherwise, to solve for the general solution v.

(c) Substitute back $u = \ln x$ and $v = x^3 y(x)$ into the solution above to obtain solution curves of the differential equation.

Problem *. In Lecture 9 we considered Hooke's Law, which describes oscillations of a spring mass without friction. Let us first study the case of *damped oscillations*, which can be described by the equation

$$mx''(t) = -kx(t) - ax'(t), \qquad x(0) = \alpha,\ x'(0) = \beta,$$

with constant coefficients and real numbers α and β. To avoid the trivial case assume $(\alpha, \beta) \neq (0, 0)$.

(a) Solve the damped oscillation equation above. You should have three different answers, corresponding to:
- *overdamping* in case $a^2 - 4km > 0$,
- *critical damping* in case $a^2 - 4km = 0$,
- *underdamping* in case $a^2 - 4km < 0$.

(b) Show that $x(t)$ crosses the origin at most once in the overdamped and critically damped case, and crosses the origin infinitely many times in the underdamped case.

Next let us study the case of *driven oscillations with damping*. This can be described by the equation

$$mx''(t) = -kx(t) - ax'(t) + f(t), \qquad x(0) = \alpha,\ x'(0) = \beta,$$

with constant coefficients and driving force $f(t)$. Here α and β can be any two chosen real numbers.

(c) Find the general formulas for the solution of this equation using Duhamel's Formula. There will be three formulas, one each for overdamping, critical damping, and underdamping.

(d) Solve

$$x''(t) = -\frac{1}{4}x(t) - x'(t) + \cos t, \qquad x(0) = 0,\ x'(0) = 0.$$

21.8 Homework 5

This homework reviews lectures 14 to 16. The last problem is optional and not for credit.

Problem 1. Solve

$$4y'''(t) + 2y''(t) + y'(t) = 1, \qquad y(0) = 1,\ y'(0) = 0,\ y''(0) = 2.$$

Problem 2. Find the general solution to

$$\left(\frac{d^2}{dt^2} + 4\right)^2 y(t) = \sin(2t).$$

Also find all solutions $y(t)$ that are bounded, in the sense that there exists a constant C_y (depending on $y(t)$) such that $|y(t)| \leq C_y$ for all $t \in \mathbb{R}$.

Problem 3. Find the general solution to

$$y''(t) - 4y'(t) + 4y(t) = \frac{2\ln(t)}{t}e^{2t}$$

on the interval $t > 0$. (This problem *will* appear in the final, so make sure you know how to do it.)

Problem 4. Find the general solution to

$$(t^2 + t)x''(t) + (2 - t^2)x'(t) - (2 + t)x(t) = t(t + 1)^2$$

by guessing two special solutions $x_1(t)$ and $x_2(t)$, and applying variation of parameters to get a particular solution $x_p(t)$. Try exponentials and polynomials for the special solutions.

Problem 5. Find the general solution to

$$4t^2y''(t) + y(t) = 24\sqrt{t}\ln t.$$

Problem 6. Find the general solution to

$$y''(t) - \tan(t)y'(t) - \sec^2(t)y(t) = \sin(t).$$

Problem 7. Let

$$M = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}, \qquad A = \begin{bmatrix} 31 & 32 \\ 32 & 40 \end{bmatrix}.$$

(a) Verify that M and A are positive definite matrices.
(b) For all values of $\omega > 0$, find the solution to

$$M\,\vec{x}''(t) = -A\,\vec{x}(t) + (3\cos(\omega t), 6\cos(\omega t)), \qquad \vec{x}(0) = \vec{x}'(0) = 0.$$

(c) Find all ω where resonance occurs.

Problem 8. Find the eigenvalues, and a corresponding orthonormal basis of eigenvectors, for the matrix

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

Problem *. The *Legendre equation* is the equation

$$(1 - x^2)y''(x) - 2xy'(x) + \alpha(\alpha + 1)y(x) = 0$$

with α a fixed constant. Let us find the general solution to this differential equation on the interval $(-1, 1)$. Suppose one has a solution of the form

$$y(x) = \sum_{n=0}^{\infty} c_n x^n.$$

(a) Compute the Maclaurin series of

$$-\frac{2x}{1 - x^2} \qquad \text{and} \qquad \frac{\alpha(\alpha + 1)}{1 - x^2}.$$

(b) Show that

$$c_{n+2} = -\frac{(\alpha+n+1)(\alpha-n)}{(n+2)(n+1)} c_n.$$

(c) Using part (b) or otherwise, show that

$$c_{2m} = (-1)^m c_0 \frac{(\alpha+2m-1)(\alpha+2m-3)\cdots(\alpha+1)\alpha(\alpha-2)\cdots(\alpha-2m+2)}{(2m)!},$$

$$c_{2m+1} = (-1)^m c_1 \frac{(\alpha+2m)(\alpha+2m-2)\cdots(\alpha+2)(\alpha-1)(\alpha-3)\cdots(\alpha-2m+1)}{(2m+1)!}.$$

(d) Show that

$$y_1(x) = 1 + \sum_{m=1}^{\infty} (-1)^m \frac{(\alpha+2m-1)\cdots(\alpha+1)\alpha\cdots(\alpha-2m+2)}{(2m)!} x^{2m}$$

$$y_2(x) = x + \sum_{m=1}^{\infty} (-1)^m \frac{(\alpha+2m)\cdots(\alpha+2)(\alpha-1)\cdots(\alpha-2m+1)}{(2m+1)!} x^{2m+1}$$

are two linearly independent solutions to the Legendre equation. You need to also check that these two solutions are convergent power series on $(-1, 1)$.

(e) Hence, or otherwise, write down the general solution to the Legendre equation.

21.9 Final Review Problems

This problem set reviews key ideas relevant to the final. The last problem is optional.

Problem 1. Solve the differential equation

$$\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t) + (12, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), \qquad \vec{\boldsymbol{x}}(1) = (0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0),$$

where A is the 8×8 matrix given by

$$\begin{bmatrix} -1 & 16 & 0 & 0 & 0 & 0 & 7 & 0 \\ -1 & 7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 2 & 5 & 9 & 0 \\ 0 & 0 & 0 & -5 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Your answer should not have integrals or terms that are not fully computed out.

Problem 2. Consider the system of differential equations

$$\vec{\boldsymbol{x}}'(t) = \begin{bmatrix} -t^{-1} & 0 \\ t^2 & t^{-1} \end{bmatrix} \vec{\boldsymbol{x}}(t) + \begin{bmatrix} 2 \\ -t^3 \end{bmatrix}.$$

Check that

$$\begin{bmatrix} t \\ t \end{bmatrix}, \qquad \begin{bmatrix} t \\ 3t \end{bmatrix}, \qquad \begin{bmatrix} t^{-1}+t \\ t^2+t \end{bmatrix},$$

are three solutions to this differential equation. Using this, find the general solution.

Problem 3. Here are two conceptual problems.

(a) If A is a 5×5 real matrix such that $(x-1)^2$ divides $\det(A - xI)$, must e^{tA} be unbounded as t approaches infinity?

(b) If A is a 5×5 real matrix with exactly three eigenvalues -2019 and $-8+i$ and $-8-i$, must every solution to $\vec{\boldsymbol{x}}'(t) = A\,\vec{\boldsymbol{x}}(t)$ be bounded as t goes to infinity? In general, how do the eigenvalues tell us about the behavior of solutions as t approaches infinity?

[Remark: Every conceptual problem in this course has involved either the characteristic polynomial or the Jordan Canonical Form of a matrix. The final will be no different.]

Problem 4. Compute the Jordan Canonical Form of

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

Problem 5. Consider the system of differential equations

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = x - x^2.$$

(a) Find the equilibrium points, and discuss the kind of stability around the equilibrium points.

(b) Show that this system is Hamiltonian, and compute the solution curve given initial conditions $x(2) = 1$ and $y(2) = 1$.

(c) For the solution curve above, find the local extrema of $y(t)$. Are there global extrema?

Problem 6. Consider the system of differential equations

$$\frac{dx}{dt} = y^2 + xy + 2y + 2x,$$
$$\frac{dy}{dt} = xy - y.$$

(a) Find the equilibrium points, and discuss the kind of stability around the equilibrium points.

(b) Show that this system is not Hamiltonian.

Problem 7. Solve the following differential equations.

(a) $y'''(t) + 3y''(t) + 3y'(t) + y(t) = e^{-t} + \cos t - 1$.

(b) $xy''(x) - (x+1)y'(x) + y(x) = 0$. (Guess exponentials and polynomials.)

(c) $y''(t) - t^{-1}y'(t) + 4t^2 y(t) = 0$. (Guess trigonometric functions.)

(d) $2t^2 y'''(t) + 3ty''(t) - 15y'(t) = 0$, with $y(1) = 1$ and $y'(1) = 2$ and $y''(1) = 1$.

(e) $y'' - 2y' + y = e^t(t^2+1)^{-1}$, with $y(0) = 2$ and $y'(0) = 5$.

Problem 8. Let

$$A = \begin{bmatrix} 3 & 4 \\ 4 & 9 \end{bmatrix}, \qquad \vec{\boldsymbol{g}}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

For all values of $\omega > 0$, find the solution to

$$\vec{\boldsymbol{x}}''(t) = -A\,\vec{\boldsymbol{x}}(t) + \cos(\omega t)\,\vec{\boldsymbol{g}}_0, \qquad \vec{\boldsymbol{x}}(0) = \vec{\boldsymbol{x}}'(0) = (1,1).$$

Also find the values of ω where resonance occurs.

Problem 9. Find the eigenvalues, and a corresponding orthonormal basis of eigenvectors, for the matrix

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

Problem 10. Find the curve that minimizes the functional

$$\int_0^1 (x'(t))^2 + (x(t))^2 \, dt$$

subject to boundary conditions $x(0) = 1$ and $x(1) = 0$.

Problem 11. Find the curve that minimizes the functional

$$\int_1^4 (x'(t)x(t))^2 \, dt$$

subject to boundary conditions $x(1) = 1$ and $x(4) = 2$.

Problem 12. Recall that the characteristic polynomial of an $n \times n$ matrix A is the polynomial

$$c(x) = \det(A - xI).$$

The Cayley-Hamilton Theorem tells us that $c(A) = 0$, and the minimal polynomial is the smallest degree polynomial of the form

$$m(x) = x^d + c_{d-1}x^{d-1} + \cdots + c_1 x + c_0$$

such that $m(A) = 0$. A problem in the midterm review tells us that the minimal polynomial divides the characteristic polynomial.

(a) Let A be a diagonalizable $n \times n$ matrix with distinct eigenvalues. Show that $m(x) = c(x)$.

(b) Let A be a general $n \times n$ matrix. Determine when $m(x) = c(x)$.

Problem *. We know from Calculus I and II that the surface of revolution of a positive function $y : [a, b] \longrightarrow \mathbb{R}$ about the x-axis is given by

$$S(y) = \int_a^b 2\pi y(x)\sqrt{1 + y'(x)} \, dx.$$

(a) Show that a function $y(x)$ that minimizes $S(y)$ must satisfy

$$y'(x) = \frac{\sqrt{(y(x))^2 - C^2}}{C}$$

for some constant C.

(b) Recall the hyperbolic cosine and hyperbolic sine functions are defined by

$$\cosh(x) = \frac{e^x + e^{-x}}{2}, \qquad \sinh(x) = \frac{e^x - e^{-x}}{2}.$$

By substituting $y(x) = C \cosh g(x)$, show that $g(x)$ satisfies

$$g'(x) = \frac{1}{C}.$$

(c) Finally, show that the only function $y(x)$ minimizing $S(y)$ is

$$y(x) = C \cosh\left(\frac{x + D}{C}\right),$$

where C and D are constants determined by $y(a)$ and $y(b)$. This curve is called the *catenary*.

21.10 Final

The final consists of eight problems and an extra credit problem.

Problem 1. Consider the system of differential equations

$$\frac{dx}{dt} = xy + x + y - 1,$$
$$\frac{dy}{dt} = x^2 - x - y.$$

Find the equilibrium points, and discuss the kind of stability around the equilibrium points.

Problem 2. Solve the differential equation

$$\vec{x}'(t) = A\,\vec{x}(t) + (12, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), \qquad \vec{x}(1) = (0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0),$$

where A is the 12×12 matrix given by

$$\begin{bmatrix} -1 & 16 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 2 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}$$

Your answer should not have integrals or terms that are not fully computed out.

Problem 3. Let

$$M = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}, \qquad A = \begin{bmatrix} 14 & 13 \\ 13 & 14 \end{bmatrix}, \qquad \vec{g}_0 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

For all values of $\omega > 0$, find the solution to

$$M\,\vec{x}''(t) = -A\,\vec{x}(t) + \cos(\omega t)\,\vec{g}_0, \qquad \vec{x}(0) = \vec{x}'(0) = (-1, 2).$$

Also find the values of ω where resonance occurs.

Problem 4. The first part is related to the second part.

(a) Using the substitution $y = \ln x$, or otherwise, compute the indefinite integrals

$$\int \sin(\ln x)\,dx \qquad \text{and} \qquad \int \cos(\ln x)\,dx.$$

(b) Solve

$$t^2 y'''(t) + t y''(t) + y'(t) = 1, \qquad y(1) = 1,\ y'(1) = 0,\ y''(1) = 2.$$

Problem 5. Let A be a 1748×1748 real matrix. Suppose A has 1 as the unique eigenvalue, and suppose A has exactly 1747 linearly independent eigenvectors.

(a) Show that

$$A^{2019} = 2019A - 2018I,$$

where I is the 1748×1748 identity matrix.

(b) Show that

$$A(2I - A) = I,$$

so A is invertible with $A^{-1} = 2I - A$.

Problem 6. Find the general solution to

$$3y''(t) - 12y'(t) + 12y(t) = \frac{6\ln(t)}{t}e^{2t}$$

on the interval $t > 0$.

Problem 7. Find the minimum value and minimizing function $y(x)$ for

$$\int_0^1 (1+x)(y'(x))^2\,dx$$

subject to the conditions $y(0) = 0$ and $y(1) = 1$.

Problem 8. I will advise you not to explicitly solve for the differential equations in this problem.

(a) Consider the differential equation

$$\vec{\boldsymbol{x}}'(t) = \begin{bmatrix} -2x(t) - 4y(t) + 5 \\ 2x(t) + 2y(t) + 1 \end{bmatrix}, \qquad \vec{\boldsymbol{x}}(0) = (1, 1).$$

Write down the equation for the graph of the solution $\vec{\boldsymbol{x}}(t)$. Hence, or otherwise, determine the minimum value of $y(t)$.

(b) Consider the solution $\vec{\boldsymbol{x}}(t) = (x(t), y(t))$ of the differential equation

$$\vec{\boldsymbol{x}}'(t) = \begin{bmatrix} 2x(t) + 2y(t) + 1 \\ 2x(t) + 4y(t) - 5 \end{bmatrix}, \qquad \vec{\boldsymbol{x}}(521) = (2019e^{\pi}, 1438\pi^{e}).$$

Compute

$$\lim_{t\to\infty} |x(t)| \qquad \text{and} \qquad \lim_{t\to\infty} |y(t)|.$$

Extra Credit. Let C be a symmetric and positive definite 2×2 real matrix, and let $h : [0, \infty) \longrightarrow \mathbb{R}$ be an integrable function. Show that

$$\int\int_{\mathbb{R}^2} h\left(\langle \vec{\boldsymbol{x}}, C\,\vec{\boldsymbol{x}} \rangle\right)\,dxdy = \frac{\pi}{\sqrt{\det(C)}} \int_0^\infty h(t)\,dt,$$

where the notation $\langle \vec{\boldsymbol{x}}, C\,\vec{\boldsymbol{x}} \rangle$ means

$$\langle \vec{\boldsymbol{x}}, C\,\vec{\boldsymbol{x}} \rangle = \begin{bmatrix} x & y \end{bmatrix} C \begin{bmatrix} x \\ y \end{bmatrix}.$$

Use this to compute

$$\int\int_{\mathbb{R}^2} \frac{1}{(1 + x^2 + 4xy + 5y^2)^2}\,dxdy.$$

21.11 Extra Credit Assignment

In this extra credit assignment we seek to understand the differential equation

$$\vec{x}''(t) = -\frac{GM}{\| \vec{x}(t)\|^3}\,\vec{x}(t)$$

with G and M constants. This is derived from Newton's Law of Gravitation and Newton's Second Law. If you have done some physics, you might realize that, together with the fact that planets have negative total energy, the above equation is the Law of Planetary Motion that governs the orbit $\vec{x}(t)$ of a planet circling around a star with mass M (and G is the gravitational constant). Our goal is to understand the above differential equation using techniques similar to understanding the solutions of first order differential equations arising from conservative vector fields. In particular, we will derive Kepler's Three Laws.

Kepler's First Law. *The orbit of the planet must be an ellipse, with the star at one of the focal points of the ellipse.*

Kepler's Second Law. *As the planet moves around the orbit, the line connecting the planet to the star sweeps out equal areas in equal time.*

Kepler's Third Law. *The square of the orbital period of the planet is proportional to the cube of the semi-major length of its orbit.*

Historically Johannes Kepler derived his laws by analyzing astronomical observations, and it was Issac Newton that realized that these laws can be derived as special consequences of his famous second law of motion. Try drawing pictures to see that Kepler's Three Laws agrees with intuition and common knowledge.

Deriving Kepler's First Law

Let m be the mass of the planet. As is tradition, let us introduce the following quantities.

- The *momentum* of the planet is defined to be

$$\mathbf{p}(t) = m\,\vec{x}'(t).$$

- The *angular momentum* of the planet is defined to be

$$\mathbf{L}(t) = \vec{x}(t) \times \mathbf{p}(t),$$

  where the cross symbol is the cross product.
- The *Runge-Lenz vector* of the planet is defined to be

$$\mathbf{A}(t) = \mathbf{p}(t) \times \mathbf{L}(t) - GMm^2\frac{\vec{x}(t)}{\| \vec{x}(t)\|}.$$

- The *total energy* of the planet is defined to be

$$\mathbf{E}(t) = \frac{\|\mathbf{p}(t)\|^2}{2m} - \frac{GMm}{\| \vec{x}(t)\|}.$$

Because of physical assumptions we always assume $\vec{x}(t)$ satisfies

$$\vec{x}''(t) = -\frac{GM}{\| \vec{x}(t)\|^3}\,\vec{x}(t),$$

and the total energy $\mathbf{E}(t)$ is negative for all t.

Problem 1. Show that
$$\frac{d}{dt}\mathbf{L}(t) = \frac{d}{dt}\mathbf{A}(t) = \frac{d}{dt}\mathbf{E}(t) = 0,$$
so that the angular momentum, Runge-Lenz vector, and total energy of the planet are actually constant vectors. Let us now denote them by $\mathbf{L}$, $\mathbf{A}$, and $\mathbf{E}$ respectively.

Problem 2. If $\mathbf{L} = 0$ then show that $\vec{\boldsymbol{x}}(t)$ must be a line. This is not an orbit, so we assume $\mathbf{L} \neq 0$ as well.

Problem 3. Show that
$$\vec{\boldsymbol{x}}(t) \cdot \mathbf{A} = \|\mathbf{L}\|^2 - GMm^2 \| \vec{\boldsymbol{x}}(t)\|$$
by showing that
$$\vec{\boldsymbol{x}}(t) \cdot (\mathbf{p}(t) \times \mathbf{L}) = \mathbf{L} \cdot \mathbf{L} = \|L\|^2.$$

Problem 4. Show that
$$\mathbf{L} \cdot \vec{\boldsymbol{x}}(t) = \mathbf{L} \cdot \mathbf{A} = 0,$$
so both $\vec{\boldsymbol{x}}(t)$ and $\mathbf{A}$ lie on the plane orthogonal to $\mathbf{L}$, called the *orbital plane.*

Problem 5. After changing coordinates on the orbital plane such that $\mathbf{A}$ is a scalar multiple of $e_1 = (1, 0)$, and writing $\vec{\boldsymbol{x}}(t) = (x(t), y(t))$, use Problem 3 to show that
$$(G^2M^2m^4 - \|\mathbf{A}\|^2)x^2(t) + G^2M^2m^4y^2(t) + 2\|\mathbf{L}\|^2\|\mathbf{A}\|x(t) = \|\mathbf{L}\|^4.$$
Notice that this is an equation of a conic section. Use this to show that $\vec{\boldsymbol{x}}(t)$ must be an ellipse by showing the following fact: our assumption that the total energy $\mathbf{E}$ is negative implies that
$$\| \vec{\boldsymbol{x}}(t)\| \leq \frac{GMm}{|\mathbf{E}|}.$$
This concludes our derivation of Kepler's First Law.

Deriving Kepler's Second Law

We now introduce some standard notation from classical geometry. Let R and r be the maximal and minimal lengths between the origin and $\vec{\boldsymbol{x}}(t)$ respectively. Then the *semi-major length* of $\vec{\boldsymbol{x}}(t)$ is defined to be
$$a = \frac{R + r}{2},$$
and the *semi-minor length* is defined to be
$$b = a\sqrt{1 - e^2}, \qquad \text{where } e = \frac{R - r}{R + r}.$$
The quantity e above is usually called the *eccentricity* of $\vec{\boldsymbol{x}}(t)$. By standard single-variable calculus we have the following assertions.

- The area of the ellipse determined by $\vec{\boldsymbol{x}}(t)$ is πab.
- There exists a number u such that
$$\left.\frac{d}{dt}\right|_u \| \vec{\boldsymbol{x}}(t)\| = 0,$$
so in particular $\vec{\boldsymbol{x}}(u)$ is perpendicular to $\vec{\boldsymbol{x}}'(u)$.

I will not ask you to show these facts, but you should convince yourself why they are true if you have not seen them before.

Problem 6. Choose u as above such that

$$\| \vec{\boldsymbol{x}}(u)\|' = \frac{d}{dt}\bigg|_u \| \vec{\boldsymbol{x}}(t)\| = 0.$$

Show that

$$\|\mathbf{L}\| = \| \vec{\boldsymbol{x}}(t)\| \|\mathbf{p}(t)\|$$

and

$$\mathbf{E}\| \vec{\boldsymbol{x}}(t)\|^2 + GMm\| \vec{\boldsymbol{x}}(t)\| - \frac{\|\mathbf{L}\|\| \vec{\boldsymbol{x}}(t)\|}{2m} = 0$$

Consequently, show that

$$\|\mathbf{A}\| = m\mathbf{E}(R - r) = GMm^2 e$$

and that the direction of A points from the origin of the orbital plane to the point of closest approach of $\vec{\boldsymbol{x}}(t)$ to the origin.

Problem 7. Let $\alpha(t)$ be the area the line connecting the star and the planet sweeps out from time 0 to time t. Show that

$$\frac{d}{dt}\alpha(t) = \frac{\|\mathbf{L}\|}{2m},$$

and explain how this can be used to derive Kepler's Second Law, by showing the following fact: the area swept out from time t to $t + \Delta$, for delta very small, is approximately a triangle with area

$$\frac{1}{2}\| \vec{\boldsymbol{x}}(t) \times \vec{\boldsymbol{x}}'(t + \Delta)\|,$$

and linear approximation tells us that this quantity is approximately

$$\frac{\Delta}{2}\| \vec{\boldsymbol{x}}(t) \times \vec{\boldsymbol{x}}'(t)\|.$$

Deriving Kepler's Third Law

Let P be the *period* of $\vec{\boldsymbol{x}}(t)$, which is the time it takes to completely sweep out the ellipse determined by $\vec{\boldsymbol{x}}(t)$ once. By Problem 7, we know that

$$\pi ab = \frac{\|\mathbf{L}\|}{2m}P.$$

Problem 8. Use Problem 6 to show that there is a real number v such that

$$\vec{\boldsymbol{x}}(v) \cdot \mathbf{A} = a(1 - e)\|\mathbf{A}\|,$$

where b is the semi-minor length. Consequently, use Problem 3 to show that

$$a(1 - e) = \frac{\|L\|^2}{GMm^2(1 + e)},$$

and use this formula and Kepler's Second Law to derive Kepler's Third Law:

$$P^2 = \frac{4\pi^2}{GM}a^3.$$

Notice that a is the radius in the case where $\vec{\boldsymbol{x}}(t)$ is a circle, reducing Kepler's Third Law to a formula which you may have seen in high school physics.