Differential Equations: An Introduction for Engineers

Matthew Charnley and Jason Nowell

July 14, 2024

Contents

Preface and About the Author(s)

A brief outline of the history of this text, and the authors involved. Also includes license information for those that may be interested in using it.

The date is the main identifier of version. The major version / edition number is raised only if there have been substantial changes.

See https://sites.rutgers.edu/matthew.charnley for more information (including contact information).

Attributions

The main inspiration for this book, as well as the vast majority of the source material, is *Notes on Diffy Qs* by Jiří Lebl . The fact that the book is freely available and open-source provided the main motivation for creating this current text. It allowed this book to be put together in a timely manner to be useful. It significantly reduced the work needed to put together a free textbook that fit the course exactly.

Introduction to this Version

This text was originally designed for the Math 244 class at Rutgers University. This class is a first course in Differential Equations for Engineering majors. This class is taken immediately after Multivariable Calculus and does not assume any knowledge of linear algebra. Prior to the design of this book, the course used Boyce and DiPrima's *Elementary Differential Equations and Boundary Value Problems* . The course provided a very brief introduction to matrices in order to get to the information necessary to handle first order systems of differential equations. With the course being redesigned to include more linear algebra, I was pointed in the direction of Jiří Lebl's *Notes on Diffy Qs* , which was meant to be a drop-in replacement for the Boyce and DiPrima text, and as of a more recent version of the text, contained an appendix on Linear Algebra.

In creating this book, I wanted to retain the style of *Notes on Diffy Qs* but shape the text into something that directly fit the course that we wanted to run. This included reorganizing some of the topics, extra contextualization of the concept of differential equations, sections devoted to modeling principles and how these equations can be derived, and guidance in using MATLAB to solve differential equations numerically. Specifically, the content added to this book is

- **Normally a reference to an Appendix goes here.** on prerequisite material to be referred to when needed. Some of the material here was pulled from Stitz and Zeager's book *Precalculus* .
- Chapter **??** contains definitions and the basics of Fourier transforms in the context of solving partial differential equations, with some information adapted from .
- Exercises and answers were added at the end of most sections of the text.

Author(s): Matthew Charnley and Jason Nowell

Acknowledgements

I would like to acknowledge David Molnar, who initially referred me to the *Notes on Diffy Qs* text , as well as the *Precalculus* text , and provided inspiration and motivation to work on designing this text. For feedback during the development of the text, I want to acknowledge David Herrera, Yi-Zhi Huang, and many others who have helped over the development and refinement of this text. Finally, I want to acknowledge the Rutgers Open and Affordable Textbook Program for supporting the development and implementation of this text.

Introduction to *Notes on Diffy Qs*

This book originated from my class notes for Math 286 at the University of Illinois at Urbana-Champaign (UIUC) in Fall 2008 and Spring 2009. It is a first course on differential equations for engineers. Using this book, I also taught Math 285 at UIUC, Math 20D at University of California, San Diego (UCSD), and Math 4233 at Oklahoma State University (OSU). Normally these courses are taught with Edwards and Penney, *Differential Equations and Boundary Value Problems: Computing and Modeling* , or Boyce and DiPrima's *Elementary Differential Equations and Boundary Value Problems* , and this book aims to be more or less a drop-in replacement. Other books I used as sources of information and inspiration are E.L. Ince's classic (and inexpensive) *Ordinary Differential Equations* , Stanley Farlow's *Differential Equations and Their Applications* , now available from Dover, Berg and McGregor's *Elementary Partial Differential Equations* , and William Trench's free book *Elementary Differential Equations with Boundary Value Problems* . See the Further Reading chapter at the end of the book.

Computer resources

The book's website `https://www.jirka.org/diffyqs/` contains the following resources:

(a) Interactive SAGE demos.

(b) Online WeBWorK homeworks (using either your own WeBWorK installation or Edfinity) for most sections, customized for this book.

(c) The PDFs of the figures used in this book.

I taught the UIUC courses using IODE (`https://faculty.math.illinois.edu/iode/`). IODE is a free software package that works with Matlab (proprietary) or Octave (free software). The graphs in the book were made with the Genius software (see `https://www.jirka.org/genius.html`). I use Genius in class to show these (and other) graphs.

Acknowledgments

Firstly, I would like to acknowledge Rick Laugesen. I used his handwritten class notes the first time I taught Math 286. My organization of this book through chapter 5, and the choice of material covered, is heavily influenced by his notes. Many examples and computations are taken from his notes. I am also heavily indebted to Rick for all the advice he has given me, not just on teaching Math 286. For spotting errors and other suggestions, I would also like to acknowledge (in no particular order): John P. D'Angelo, Sean Raleigh, Jessica Robinson, Michael Angelini, Leonardo Gomes, Jeff Winegar, Ian Simon, Thomas Wicklund, Eliot Brenner, Sean Robinson, Jannett Susberry, Dana Al-Quadi, Cesar Alvarez, Cem Bagdatlioglu, Nathan Wong, Alison Shive, Shawn White, Wing Yip Ho, Joanne Shin, Gladys Cruz, Jonathan Gomez, Janelle Louie, Navid Froutan, Grace Victorine, Paul Pearson, Jared Teague, Ziad Adwan, Martin Weilandt, Sönmez Sahutoglu,Pete Peterson, Thomas Gresham, Prentiss Hyde, Jai Welch, Simon Tse, Andrew Browning, James Choi, Dusty Grundmeier, John Marriott, Jim Kruidenier, Barry Conrad, Wesley Snider, Colton Koop, Sarah Morse, Erik Boczko, Asif Shakeel, Chris Peterson, Nicholas Hu, Paul Seeburger, Jonathan McCormick, David Leep, William Meisel, Shishir Agrawal, Tom Wan, Andres Valloud, and probably others I have forgotten. Finally, I would like to acknowledge NSF grants DMS-0900885 and DMS-1362337.

Introduction to differential equations

We Introduction some basics of differential equations

Differential equations

Consider the following situation:

> An object falling through the air has its velocity affected by two factors: gravity and a drag force. The velocity downward is increased at a rate of 9.8 m/s^2 due to gravity, and it is decreased by a rate equation to 0.3 times the current velocity of the object. If the ball is initially thrown downwards at a speed of 2 m/s, what will the velocity be 10 seconds later?

There might be enough information here to determine the velocity at any later point in time (it turns out, there is) but the information given isn't really about the velocity. Rather, information is given about the rate of change of the velocity. We know that the velocity will be increased at a rate of $9.8m/s^2$ due to gravity. How can this be interpreted? The rate of change has been discussed previously way back in Calculus 1; this is the derivative. Thus, if we let the unknown function $v(t)$ represent the velocity of the object, the description above gives information about the derivative of this function for$v(t)$. Taking the two different factors (the increase and decrease of velocity) into account, we can write an expression for this derivative, giving that

$$\frac{dv}{dt} = 9.8 - 0.3v.$$

Even though we don't know what $v(t)$ is, we know that it must affect the derivative of the velocity in this particular way, so we can write this equation. That's why we give a name to this function, so that we can use it in writing this question, which, since it is an equation involving the derivative of an unknown function $v(t)$, we call this a differential equation. Our goal here would be to use this information, plus the fact that the velocity at time zero is $v(0) = 2$ m/s to find the value of $v(10)$, or, more generally, the function $v(t)$ for any t.

The laws of physics, beyond just that of simple velocity, are generally written down as differential equations. Therefore, all of science and engineering use differential equations to some degree. Understanding differential equations is essential to understanding almost anything you will study in your science and engineering classes. You can think of mathematics as the language of science, and differential equations are one of the most important parts of this language as far as science and engineering are concerned. As an analogy, suppose all your classes from now on were given half in Swahili and half in English. It would be important to first learn Swahili, or you would have a very tough time getting a good grade in your classes. Without it, you might be able to make sense of some of what is going on, but would definitely be missing an important part of the picture.

Definition 1. *A* differential equation *is an equation that involves one or more derivatives of an unknown function. For a differential equation, the* order *of the differential equation is the highest order derivative that appears in the equation.*

One example of a first order differential equation is

$$\frac{dx}{dt} + x = 2\cos t. \tag{1}$$

Here x is the *dependent variable* and t is the *independent variable*. Note that we can use any letter we want for the dependent and independent variables. This equation arises from Newton's law of cooling where the ambient temperature oscillates with time.

To make sure that everything is well-defined, we will assume that we can always write our differential equation with the highest order derivative written as a function of all lower derivatives and the independent variable. For the previous example, since we can write (1) as

$$\frac{dx}{dt} = 2\cos t - x$$

Learning outcomes: Identify a differential equation and determine the order of a differential equation Verify that a function is a solution to a differential equation Solve some fundamental differential equations.

Author(s): Matthew Charnley and Jason Nowell

where the highest derivative x' is written as a function of t and x, we have a proper differential equation. On the other hand, something like

$$\left(\frac{dy}{dt}\right)^2 + y^2 = 1 \tag{2}$$

is not a proper differential equation because we can't solve for $\frac{dy}{dt}$. This expression could either be written as

$$\frac{dy}{dt} = \sqrt{1-y^2} \qquad \text{or} \qquad \frac{dy}{dt} = -\sqrt{1-y^2},$$

and while both of these are proper differential equations, the version in (2) is not.

For some equations, like $y' = y^2$, the independent variable is not explicitly stated. We could be looking for a function $y(t)$ or a function $y(x)$ (or y of any other variable) and without any other information, any of these is correct. Usually, there will be information in the problem statement to indicate that the independent variable is something like time, in which case everything should be written in terms of t. It is for this reason that Leibniz notation is preferred for derivatives; an equation like

$$\frac{dy}{dt} = y^2$$

is unambiguously looking for any answer $y(t)$.

Example 1. *All of the below are differential equations*

$$\frac{dy}{dt} = e^t y \qquad\qquad z'' + z^2 = t \sin z$$

$$\frac{d^4 f}{dx^4} - 3x\frac{d^2 f}{dx^2} = x \qquad\qquad y''' + (y'')^2 - 3y = t^4.$$

Note that any letter can be used for the unknown function and its dependent variable. From the context of the equations, we can see that the unknown functions we are looking for in these examples are $y(t)$, $z(t)$, $y(x)$, and $y(t)$ respectively. The order of these equations are 1, 2, 4, and 3 respectively.

Solutions of differential equations

Solving the differential equation means finding the function that, when we plug it into the differential equation, gives a true statement. For example, take (1) from the previous section. In this case, this means that we want to find a function of t, which we call x, such that when we plug x, t, and $\frac{dx}{dt}$ into (1), the equation holds; that is, the left hand side equals the right hand side. It is the same idea as it would be for a normal (algebraic) equation of just x and t. We claim that

$$x = x(t) = \cos t + \sin t$$

is a *solution*. How do we check? We simply plug x into equation (1)! First we need to compute $\frac{dx}{dt}$. We find that $\frac{dx}{dt} = -\sin t + \cos t$. Now let us compute the left-hand side of (1).

$$\frac{dx}{dt} + x = \underbrace{(-\sin t + \cos t)}_{\frac{dx}{dt}} + \underbrace{(\cos t + \sin t)}_{x} = 2\cos t.$$

Yay! We got precisely the right-hand side. But there is more! We claim $x = \cos t + \sin t + e^{-t}$ is also a solution. Let us try,

$$\frac{dx}{dt} = -\sin t + \cos t - e^{-t}.$$

We plug into the left-hand side of (1)

$$\frac{dx}{dt} + x = \underbrace{(-\sin t + \cos t - e^{-t})}_{\frac{dx}{dt}} + \underbrace{(\cos t + \sin t + e^{-t})}_{x} = 2\cos t.$$

And it works yet again!

So there can be many different solutions. For this equation all solutions can be written in the form

$$x = \cos t + \sin t + Ce^{-t},$$

for some constant C. Different constants C will give different solutions, so there are really infinitely many possible solutions. See **Normally a reference to a previous figure goes here.** for the graph of a few of these solutions. We do not yet know how to find this solution, but we will get to that in the next chapter.

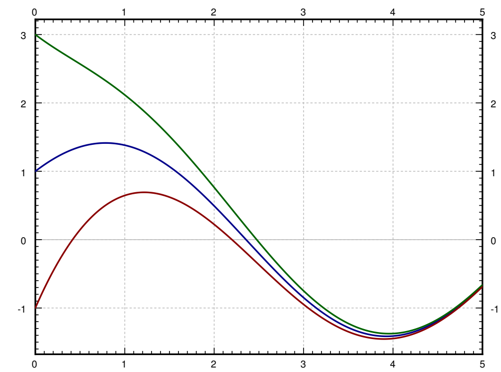


Figure 1: Few solutions of $\frac{dx}{dt} + x = 2\cos t$.

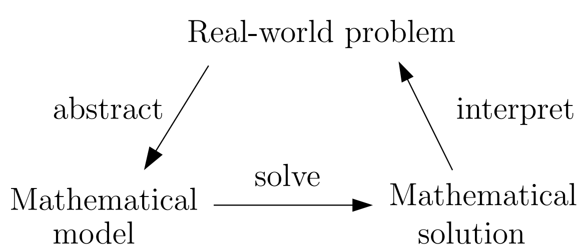
Solving differential equations can be quite hard. There is no general method that solves every differential equation. We will generally focus on how to get exact formulas for solutions of certain differential equations, but we will also spend a little bit of time on getting approximate solutions. And we will spend some time on understanding the equations without solving them.

Most of this book is dedicated to *ordinary differential equations* or ODEs, that is, equations with only one independent variable, where derivatives are only with respect to this one variable. If there are several independent variables, we get *partial differential equations* or PDEs.

Even for ODEs, which are very well understood, it is not a simple question of turning a crank to get answers. When you can find exact solutions, they are usually preferable to approximate solutions. It is important to understand how such solutions are found. Although in real applications you will leave much of the actual calculations to computers, you need to understand what they are doing. It is often necessary to simplify or transform your equations into something that a computer can understand and solve. You may even need to make certain assumptions and changes in your model to achieve this.

To be a successful engineer or scientist, you will be required to solve problems in your job that you have never seen before. It is important to learn problem solving techniques, so that you may apply those techniques to new problems. A common mistake is to expect to learn some prescription for solving all the problems you will encounter in your later career. This course is no exception.

Differential equations in practice



So how do we use differential equations in science and engineering? The main way this takes place is through the process of mathematical modeling. First, we have some *real-world problem* we wish to understand. We make some simplifying assumptions and create a *mathematical model*, which is a translation of this real-world problem into a set of differential equations. Think back to the example at the beginning of this section. We took a physical situation (a falling object) with some knowledge about how it behaves and turned that into a differential equation that describes the velocity over time. Then we apply mathematics to get some sort of a *mathematical solution*. Finally, we need to interpret our results, determining what this mathematical solution says about the real-world problem we started with. For instance, in the example at the start of the section, we could find the function $v(t)$, but then need to interpret that if we were to plug 10 into this function, we will get the velocity 10 seconds later.

Learning how to formulate the mathematical model and how to interpret the results is what your physics and engineering classes do. In this course, we will focus mostly on the mathematical analysis. This will be interspersed with discussions of this modeling process to give some context to what we are doing, and give practice for what will be seen in future physics and engineering classes.

Let us look at an example of this process. One of the most basic differential equations is the standard *exponential growth model*. Let P denote the population of some bacteria on a Petri dish. We assume that there is enough food and enough space. Then the rate of growth of bacteria is proportional to the population—a large population grows quicker. Let t denote time (say in seconds) and P the population. Our model is

$$\frac{dP}{dt} = kP,$$

for some positive constant $k > 0$.

Example 2. *Suppose there are 100 bacteria at time 0 and 200 bacteria 10 seconds later. How many bacteria will there be 1 minute from time 0 (in 60 seconds)?*

Solution: First we need to solve the equation. We claim that a solution is given by

$$P(t) = Ce^{kt},$$

where C is a constant. Let us try:

$$\frac{dP}{dt} = Cke^{kt} = kP.$$

And it really is a solution.

OK, now what? We do not know C, and we do not know k. But we know something. We know $P(0) = 100$, and we know $P(10) = 200$. Let us plug these conditions in and see what happens.

$$100 = P(0) = Ce^{k0} = C,$$
$$200 = P(10) = 100\, e^{k10}.$$

Therefore, $2 = e^{10k}$ or $\frac{\ln 2}{10} = k \approx 0.069$. So

$$P(t) = 100\, e^{(\ln 2)t/10} \approx 100\, e^{0.069t}.$$

At one minute, $t = 60$, the population is $P(60) = 6400$. See **Normally a reference to a previous figure goes here.**.

Let us talk about the interpretation of the results. Does our solution mean that there must be exactly 6400 bacteria on the plate at 60s? No! We made assumptions that might not be true exactly, just approximately. If our assumptions are reasonable, then there will be approximately 6400 bacteria. Also, in real life P is a discrete quantity, not a real number. However, our model has no problem saying that for example at 61 seconds, $P(61) \approx 6859.35$.

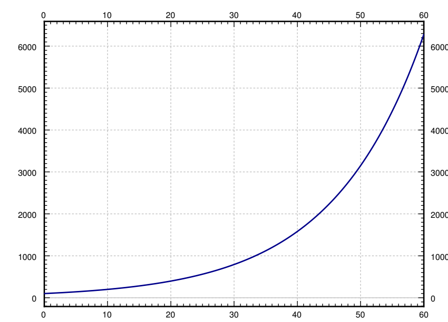


Figure 2: Bacteria growth in the first 60 seconds.

Normally, the k in $P' = kP$ is known, and we want to solve the equation for different *initial conditions*. What does that mean? Take $k = 1$ for simplicity. Suppose we want to solve the equation $\frac{dP}{dt} = P$ subject to $P(0) = 1000$ (the initial condition). Then the solution turns out to be (exercise)

$$P(t) = 1000\, e^{t}.$$

We call $P(t) = Ce^t$ *the general solution*, as every solution of the equation can be written in this form for some constant C. We need an initial condition to find out what C is, in order to find the *particular solution* we are looking for. Generally, when we say "particular solution", we just mean some solution.

Four fundamental equations

A few equations appear often and it is useful to know what their solutions are. Let us call them the four fundamental equations. Their solutions are reasonably easy to guess by recalling properties of exponentials, sines, and cosines. They are also simple to check, which is something that you should always do. No need to wonder if you remembered the solution correctly. It is good to have these as solutions that you "know" to build from when we learn solutions to other differential equations later on. In Chapter **??** we will cover the first two, and the last two will be discussed in Chapter **??**.

First such equation is

$$\frac{dy}{dx} = ky,$$

for some constant $k > 0$. Here y is the dependent and x the independent variable. The general solution for this equation is

$$y(x) = Ce^{kx}.$$

We saw above that this function is a solution, although we used different variable names.

Next,

$$\frac{dy}{dx} = -ky,$$

for some constant $k > 0$. The general solution for this equation is

$$y(x) = Ce^{-kx}.$$

Exercise 1 *Check that the y given is really a solution to the equation.*

Next, take the *second order differential equation*

$$\frac{d^2y}{dx^2} = -k^2y,$$

for some constant $k > 0$. The general solution for this equation is

$$y(x) = C_1 \cos(kx) + C_2 \sin(kx).$$

Since the equation is a second order differential equation, we have two constants in our general solution.

Exercise 2 *Check that the y given is really a solution to the equation.*

Finally, consider the second order differential equation

$$\frac{d^2y}{dx^2} = k^2y,$$

for some constant $k > 0$. The general solution for this equation is

$$y(x) = C_1e^{kx} + C_2e^{-kx},$$

or

$$y(x) = D_1 \cosh(kx) + D_2 \sinh(kx).$$

For those that do not know, cosh and sinh are defined by

$$\cosh x = \frac{e^x + e^{-x}}{2}, \qquad \sinh x = \frac{e^x - e^{-x}}{2}.$$

They are called the *hyperbolic cosine* and *hyperbolic sine*. These functions are sometimes easier to work with than exponentials. They have some nice familiar properties such as $\cosh 0 = 1$, $\sinh 0 = 0$, and $\frac{d}{dx}\cosh x = \sinh x$ (no that is not a typo) and $\frac{d}{dx}\sinh x = \cosh x$.

Exercise 3 *Check that both forms of the y given are really solutions to the equation.*

Example 3. *In equations of higher order, you get more constants you must solve for to get a particular solution. The equation $\frac{d^2y}{dx^2} = 0$ has the general solution $y = C_1x + C_2$; simply integrate twice and don't forget about the constant of integration. Consider the initial conditions $y(0) = 2$ and $y'(0) = 3$. We plug in our general solution and solve for the constants:*

$$2 = y(0) = C_1 \cdot 0 + C_2 = C_2, \qquad 3 = y'(0) = C_1.$$

In other words, $y = 3x + 2$ is the particular solution we seek.

`intro/practice/intro-practice1.tex`

Practice for Solving Exact ODEs

Why?

*Note: Exercises marked with * have answers in the back of the book.*

Exercise 4 *Show that* $x = e^{4t}$ *is a solution to* $x''' - 12x'' + 48x' - 64x = 0$.

Exercise 5 *Show that* $x = e^{-2t}$ *is a solution to* $x'' + 4x' + 4x = 0$.

Exercise 6 *Show that* $x = e^{t}$ *is not a solution to* $x''' - 12x'' + 48x' - 64x = 0$.

Exercise 7 *Is* $y = \sin t$ *a solution to* $\left(\frac{dy}{dt}\right)^2 = 1 - y^2$*? Justify.*

Exercise 8 *Is* $y = x^2$ *a solution to* $x^2y'' - 2y = 0$*? Justify.*

Exercise 9 *Let* $y'' + 2y' - 8y = 0$*. Now try a solution of the form* $y = e^{rx}$ *for some (unknown) constant* r*. Is this a solution for some* r*? If so, find all such* r*.*

Exercise 10 *Let* $xy'' - y' = 0$*. Try a solution of the form* $y = x^r$*. Is this a solution for some* r*? If so, find all such* r*.*

Exercise 11 *Verify that* $x = Ce^{-2t}$ *is a solution to* $x' = -2x$*. Find* C *to solve for the initial condition* $x(0) = 100$*.*

Exercise 12 *Verify that* $x = C_1e^{-t} + C_2e^{2t}$ *is a solution to* $x'' - x' - 2x = 0$*. Find* C_1 *and* C_2 *to solve for the initial conditions* $x(0) = 10$ *and* $x'(0) = 0$*.*

Exercise 13 *Verify that* $x = C_1e^{t} + C_2$ *is a solution to* $x'' - x' = 0$*. Find* C_1 *and* C_2 *so that* x *satisfies* $x(0) = 10$ *and* $x'(0) = 100$*.*

Exercise 14 *Find a solution to* $(x')^2 + x^2 = 4$ *using your knowledge of derivatives of functions that you know from basic calculus.*

Exercise 15 *Solve* $\frac{d\varphi}{ds} = 8\varphi$ *and* $\varphi(0) = -9$*.*

Author(s): Matthew Charnley and Jason Nowell

Exercise 16 *Solve:*

a) $\dfrac{dA}{dt} = -10A, \quad A(0) = 5$

b) $\dfrac{dH}{dx} = 3H, \quad H(0) = 1$

c) $\dfrac{d^2y}{dx^2} = 4y, \quad y(0) = 0, \quad y'(0) = 1$

d) $\dfrac{d^2x}{dy^2} = -9x, \quad x(0) = 1, \quad x'(0) = 0$

Exercise 17 *Solve:*

a) $\dfrac{dx}{dt} = -4x, \quad x(0) = 9$

b) $\dfrac{d^2x}{dt^2} = -4x, \quad x(0) = 1, \quad x'(0) = 2$

c) $\dfrac{dp}{dq} = 3p, \quad p(0) = 4$

d) $\dfrac{d^2T}{dx^2} = 4T, \quad T(0) = 0, \quad T'(0) = 6$

Exercise 18 *Is there a solution to $y' = y$, such that $y(0) = y(1)$?*

Exercise 19 *The population of city X was 100 thousand 20 years ago, and the population of city X was 120 thousand 10 years ago. Assuming constant growth, you can use the exponential population model (like for the bacteria). What do you estimate the population is now?*

Exercise 20 *Suppose that a football coach gets a salary of one million dollars now, and a raise of 10% every year (so exponential model, like population of bacteria). Let s be the salary in millions of dollars, and t is time in years.*

a) *What is $s(0)$ and $s(1)$.*

b) *Approximately how many years will it take for the salary to be 10 million.*

c) *Approximately how many years will it take for the salary to be 20 million.*

d) *Approximately how many years will it take for the salary to be 30 million.*

Classification of differential equations

We introduce how to classify various properties of differential equations.

There are many types of differential equations, and we classify them into different categories based on their properties. Let us quickly go over the most basic classification. We already saw the distinction between ordinary and partial differential equations:

Definition 2.

- Ordinary differential equations *or (ODE) are equations where the derivatives are taken with respect to only one variable. That is, there is only one independent variable.*
- Partial differential equations *or (PDE) are equations that depend on partial derivatives of several variables. That is, there are several independent variables.*

Let us see some examples of ordinary differential equations:

$$\begin{aligned}
&\frac{dy}{dt} = ky, && \text{(Exponential growth)}\\
&\frac{dy}{dt} = k(A-y), && \text{(Newton's law of cooling)}\\
&m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + kx = f(t). && \text{(Mechanical vibrations)}
\end{aligned}$$

And of partial differential equations:

$$\begin{aligned}
&\frac{\partial y}{\partial t} + c\frac{\partial y}{\partial x} = 0, && \text{(Transport equation)}\\
&\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, && \text{(Heat equation)}\\
&\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}. && \text{(Wave equation in 2 dimensions)}
\end{aligned}$$

If there are several equations working together, we have a so-called *system of differential equations*. For example,

$$y' = x, \qquad x' = y$$

is a simple system of ordinary differential equations. Maxwell's equations for electromagnetics,

$$\begin{aligned}
&\nabla \cdot \vec{D} = \rho, && \nabla \cdot \vec{B} = 0,\\
&\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, && \nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t},
\end{aligned}$$

are a system of partial differential equations. The divergence operator $\nabla\cdot$ and the curl operator $\nabla\times$ can be written out in partial derivatives of the functions involved in the x, y, and z variables.

In the first chapter, we will start attacking first order ordinary differential equations, that is, equations of the form $\frac{dy}{dx} = f(x,y)$. In general, lower order equations are easier to work with and have simpler behavior, which is why we start with them.

We also distinguish how the dependent variables appear in the equation (or system).

Definition 3. *We say an equation is* linear *if the dependent variable (or variables) and their derivatives appear linearly, that is only as first powers, they are not multiplied together, and no other functions of the dependent variables appear. Otherwise, the equation is called* nonlinear*.*

Learning outcomes: Classify equation as ordinary or partial differential equations Identify whether an equation is linear or non-linear Classify linear equations as homogeneous, non-homogeneous, or constant coefficient, as appropriate

Author(s): Matthew Charnley and Jason Nowell

Another way to determine if a differential equation is linear is if the equation is a sum of terms, where each term is some function of the independent variables or some function of the independent variables multiplied by a dependent variable or its derivative. That is, an ordinary differential equation is linear if it can be put into the form

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = b(x). \tag{3}$$

The functions a_0, a_1, ..., a_n are called the *coefficients*. The equation is allowed to depend arbitrarily on the independent variable. So

$$e^x \frac{d^2y}{dx^2} + \sin(x)\frac{dy}{dx} + x^2 y = \frac{1}{x} \tag{4}$$

is still a linear equation as y and its derivatives only appear linearly. The equation

$$\cos(x)\frac{d^2y}{dx^2} - xy + \frac{e^x}{x} = 0$$

is also linear, even though it is not initially in the correct form. From this equation, we can move the last term over to the right-hand side as a $-\frac{e^x}{x}$, and then it is in the correct form, with the $\frac{dy}{dx}$ term missing (or has coefficient zero).

All the equations and systems above as examples are linear. It may not be immediately obvious for Maxwell's equations unless you write out the divergence and curl in terms of partial derivatives. Let us see some nonlinear equations. For example Burger's equation,

$$\frac{\partial y}{\partial t} + y\frac{\partial y}{\partial x} = \nu \frac{\partial^2 y}{\partial x^2},$$

is a nonlinear second order partial differential equation. It is nonlinear because y and $\frac{\partial y}{\partial x}$ are multiplied together. The equation

$$\frac{dx}{dt} = x^2 \tag{5}$$

is a nonlinear first order differential equation as there is a second power of the dependent variable x.

Definition 4. *A linear equation may further be called* homogeneous *if all terms depend on the dependent variable. That is, if no term is a function of the independent variables alone. Otherwise, the equation is called* nonhomogeneous *or* inhomogeneous*.*

For example, the exponential growth equation, the wave equation, or the transport equation above are homogeneous. The mechanical vibrations equation above is nonhomogeneous as long as $f(t)$ is not the zero function. Similarly, if the ambient temperature A is nonzero, Newton's law of cooling is nonhomogeneous. A homogeneous linear ODE can be put into the form

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = 0.$$

Compare to (3) and notice there is no function $b(x)$.

If the coefficients of a linear equation are actually constant functions, then the equation is said to have *constant coefficients*. The coefficients are the functions multiplying the dependent variable(s) or one of its derivatives, not the function $b(x)$ standing alone. A constant coefficient nonhomogeneous ODE is an equation of the form

$$a_n\frac{d^n y}{dx^n} + a_{n-1}\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1\frac{dy}{dx} + a_0 y = b(x),$$

where $a_0, a_1, \ldots, a_n$ are all constants, but b may depend on the independent variable x. The mechanical vibrations equation above is a constant coefficient nonhomogeneous second order ODE. The same nomenclature applies to PDEs, so the transport equation, heat equation and wave equation are all examples of constant coefficient linear PDEs.

Finally, an equation (or system) is called *autonomous* if the equation does not explicitly depend on the independent variable. For autonomous ordinary differential equations, the independent variable is then thought of as time. Autonomous equation means an equation that does not change with time. For example, Newton's law of cooling is autonomous, so is equation (5). On the other hand, mechanical vibrations or (4) are not autonomous.

`intro/practice/classification-practice1.tex`

Practice for Classifying Equations

Why?

Exercise 21 *Classify the following equations. Are they ODE or PDE? Is it an equation or a system? What is the order? Is it linear or nonlinear, and if it is linear, is it homogeneous, constant coefficient? If it is an ODE, is it autonomous?*

a) $\sin(t)\dfrac{d^2x}{dt^2} + \cos(t)x = t^2$

b) $\dfrac{\partial u}{\partial x} + 3\dfrac{\partial u}{\partial y} = xy$

c) $y'' + 3y + 5x = 0, \quad x'' + x - y = 0$

d) $\dfrac{\partial^2 u}{\partial t^2} + u\dfrac{\partial^2 u}{\partial s^2} = 0$

e) $x'' + tx^2 = t$

f) $\dfrac{d^4x}{dt^4} = 0$

Exercise 22 *Classify the following equations. Are they ODE or PDE? Is it an equation or a system? What is the order? Is it linear or nonlinear, and if it is linear, is it homogeneous, constant coefficient? If it is an ODE, is it autonomous?*

a) $\dfrac{\partial^2 v}{\partial x^2} + 3\dfrac{\partial^2 v}{\partial y^2} = \sin(x)$

b) $\dfrac{dx}{dt} + \cos(t)x = t^2 + t + 1$

c) $\dfrac{d^7F}{dx^7} = 3F(x)$

d) $y'' + 8y' = 1$

e) $x'' + tyx' = 0, \quad y'' + txy = 0$

f) $\dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial s^2} + u^2$

Exercise 23 *If $\vec{u} = (u_1, u_2, u_3)$ is a vector, we have the divergence $\nabla \cdot \vec{u} = \dfrac{\partial u_1}{\partial x} + \dfrac{\partial u_2}{\partial y} + \dfrac{\partial u_3}{\partial z}$ and curl $\nabla \times \vec{u} = \left(\dfrac{\partial u_3}{\partial y} - \dfrac{\partial u_2}{\partial z}, \dfrac{\partial u_1}{\partial z} - \dfrac{\partial u_3}{\partial x}, \dfrac{\partial u_2}{\partial x} - \dfrac{\partial u_1}{\partial y}\right)$. Notice that curl of a vector is still a vector. Write out Maxwell's equations in terms of partial derivatives and classify the system.*

Exercise 24 *Suppose F is a linear function, that is, $F(x, y) = ax + by$ for constants a and b. What is the classification of equations of the form $F(y', y) = 0$.*

Exercise 25 *Write down an explicit example of a third order, linear, nonconstant coefficient, nonautonomous, nonhomogeneous system of two ODE such that every derivative that could appear, does appear.*

Exercise 26 *Write down the general zeroth order linear ordinary differential equation. Write down the general solution.*

Exercise 27 *For which k is $\dfrac{dx}{dt} + x^k = t^{k+2}$ linear. Hint: there are two answers.*

Author(s): Matthew Charnley and Jason Nowell

Exercise 28 *Write out an explicit example of a non-homogeneous fourth order, linear, constant coefficient differential equation. where all possible derivatives of the unknown function y appear.*

Exercise 29 *Let x, y, and z be three functions of t defined by the system of differential equations*

$$x' = y \quad y' = z \quad z' = 3x - 2y + 5z + e^t$$

with initial conditions $x(0) = 3$, $y(0) = -2$ and $z(0) = 1$, and let $u(t)$ be the function defined by the solution to

$$u''' - 5u'' + 2u' - 3u = e^t$$

with initial conditions $u(0) = 3$, $u'(0) = -2$, and $u''(0) = 1$.

a) *Use the substitution $u = x$, $u' = y$, and $u'' = z$ to verify that $x(t) = u(t)$ because they solve the same initial value problem.*

b) *What is the order of the system defining x, y, and z and how many components does it have?*

c) *What is the order of the equation defining u? How many components does that have?*

d) *How do these numbers relate to each other?*

Integrals as solutions

Stuff about Integrals as solutions

A first order ODE is an equation of the form

$$\frac{dy}{dx} = f(x, y),$$

or just

$$y' = f(x, y).$$

Some examples that fit this form are

$$y' = x^2 y - e^x \sin x$$

and

$$y' = e^y (x^2 + 1) - \cos(y).$$

Looking back at the last section, the first of these is linear and the second is not. In general, there is no simple formula or procedure one can follow to find solutions. In the next few sections we will look at special cases where solutions are not difficult to obtain. In this section, let us assume that f is a function of x alone, that is, the equation is

$$y' = f(x). \tag{6}$$

We could just integrate (antidifferentiate) both sides with respect to x.

$$\int y'(x)\, dx = \int f(x)\, dx + C,$$

that is

$$y(x) = \int f(x)\, dx + C.$$

This $y(x)$ is actually the general solution. So to solve (6), we find some antiderivative of $f(x)$ and then we add an arbitrary constant to get the general solution.

Now is a good time to discuss a point about calculus notation and terminology. One of the final keystone concepts in Calculus 1 is that of the fundamental theorem of calculus, which ties together two mathematical ideas: definite integrals (defined as the area under a curve) and indefinite integrals or antidifferentiation (undoing the operation of differentiation). This theorem says that these two ideas are in some sense the same; in order to compute a definite integral, one can first find an antiderivative and plug in the endpoints (the most common use of the theorem), and that the derivative of a definite integral gives back the function inside (something that will be useful in this course).

The main distinction between these two is the type of object that they are. Definite integrals evaluate to numbers, so they are functions, which means they are the object we want to deal with in this course. Indefinite integrals are families of functions, and while they have their uses (motivating the idea of a general solution), their main use is the process of antidifferentiation which leads us to solutions in the form of definite integrals. Provided that you can evaluate the antiderivative in question, these two processes will end up at exactly the same solution. If you end up confused about the terminology, the goal for this class is always a definite integral, but we can use antiderivatives to get there. Hence the terminology *to integrate* when we may really mean *to antidifferentiate*. Integration is just one way to compute the antiderivative (and it is a way that always works, see the following examples). Integration is defined as the area under the graph and it also happens to also compute antiderivatives. For sake of consistency, we will keep using the indefinite integral notation when we want an antiderivative, and you should *always* think of the definite integral as a way to write it.

Example 4. *Find the general solution of $y' = 3x^2$.*

Solution: Elementary calculus tells us that the general solution must be $y = x^3 + C$. Let us check by differentiating: $y' = 3x^2$. We got *precisely* our equation back.

Learning outcomes: Solve a first order differential equation by direct integration Understand the difference between a general solution and the solution to an initial value problem.

Author(s): Matthew Charnley and Jason Nowell

Normally, we will also have an initial condition such as $y(x_0) = y_0$ for some two numbers x_0 and y_0 (x_0 is often 0, but not always). If we do, the combination of a differential equation and an initial condition is called an initial value problem. We can then write the solution as a definite integral in a nice way. Suppose our problem is $y' = f(x)$, $y(x_0) = y_0$. Then the solution is

$$y(x) = \int_{x_0}^{x} f(s)\, ds + y_0. \tag{7}$$

Let us check! We compute

$$y'(x) = \frac{d}{dx}\left[\int_{x_0}^{x} f(s)\, ds + y_0\right].$$

Since y_0 is a constant, it's derivative is zero, and by the fundamental theorem of calculus

$$\frac{d}{dx}\int_{x_0}^{x} f(s)\ dx = f(x).$$

Therefore $y' = f(x)$, and by Jupiter, y is a solution. Is it the one satisfying the initial condition? Well,

$$y(x_0) = \int_{x_0}^{x_0} f(x)\, dx + y_0$$

and since f is a nice function, we know that the integral of f with matching endpoints is 0. Therefore $y(x_0) = y_0$. It is!

Do note that the definite integral and the indefinite integral (antidifferentiation) are completely different beasts. The definite integral always evaluates to a number. Therefore, (7) is a formula we can plug into the calculator or a computer, and it will be happy to calculate specific values for us. We will easily be able to plot the solution and work with it just like with any other function. It is not so crucial to always find a closed form for the antiderivative.

Example 5. *Solve*

$$y' = e^{-x^2}, \qquad y(0) = 1.$$

Solution: By the preceding discussion, the solution must be

$$y(x) = \int_{0}^{x} e^{-s^2}\ ds + 1.$$

Here is a good way to make fun of your friends taking second semester calculus. Tell them to find the closed form solution. Ha ha ha (bad math joke). It is not possible (in closed form). There is absolutely nothing wrong with writing the solution as a definite integral. This particular integral is in fact very important in statistics.

While there is nothing wrong with writing solutions as a definite integral, they should be simplified and evaluated if possible. Given the differential equation

$$y' = 3x^2, \qquad y(2) = 6,$$

the solution can be written as

$$y(x) = \int_{2}^{x} 3s^2\ ds + 6.$$

However, it is much more convenient, both for human reasoning and computers, to write this solution as

$$y(x) = x^3 - 2.$$

So, if integrals can be evaluated and simplified to explicit functions, then they should be worked out. If it is not possible, then answers in integral form are completely fine.

Classical problems leading to differential equations solvable by integration are problems dealing with velocity, acceleration and distance. You have surely seen these problems before in your calculus class.

Example 6. *Suppose a car drives at a speed $e^{t/2}$ meters per second, where t is time in seconds. How far did the car get in 2 seconds (starting at $t = 0$)? How far in 10 seconds?*

Solution: Let x denote the distance the car traveled. The equation is

$$x' = e^{t/2}.$$

We just integrate this equation to get that

$$x(t) = 2e^{t/2} + C.$$

We still need to figure out C. We know that when $t = 0$, then $x = 0$. That is, $x(0) = 0$. So

$$0 = x(0) = 2e^{0/2} + C = 2 + C.$$

Thus $C = -2$ and

$$x(t) = 2e^{t/2} - 2.$$

Now we just plug in to get where the car is at 2 and at 10 seconds. We obtain

$$x(2) = 2e^{2/2} - 2 \approx 3.44 \text{ meters}, \qquad x(10) = 2e^{10/2} - 2 \approx 294 \text{ meters}.$$

Example 7. *Suppose that the car accelerates at a rate of t^2 m/s². At time $t = 0$ the car is at the 1 meter mark and is traveling at 10 m/s. Where is the car at time $t = 10$?*

Solution: Well this is actually a second order problem. If x is the distance traveled, then x' is the velocity, and x'' is the acceleration. The initial value problem for this situation is

$$x'' = t^2, \qquad x(0) = 1, \qquad x'(0) = 10.$$

What if we say $x' = v$. Then we have the problem

$$v' = t^2, \qquad v(0) = 10.$$

Once we solve for v, we can integrate and find x.

Exercise 30 *Solve for v, and then solve for x. Find $x(10)$ to answer the question.*

`firstOrder/practice/directIntegration-practice1.tex`

Practice for Solving by Direct Integration

Why?

Exercise 31 *Solve* $\dfrac{dy}{dx} = x^2 + x$ *with* $y(1) = 3$.

Exercise 32 *Solve* $\dfrac{dy}{dx} = \sin(5x)$ *with* $y(0) = 2$.

Exercise 33 *Solve* $\dfrac{dy}{dx} = e^x + x$ *with* $y(0) = 10$.

Exercise 34 *Solve* $\dfrac{dy}{dx} = 2xe^{3x}$ *with* $y(0) = 1$.

Exercise 35 *Solve* $\dfrac{dx}{dt} = e^t \cos(2t) + t$ *with* $x(0) = 3$.

Exercise 36 *Solve* $\dfrac{dy}{dx} = \dfrac{1}{x^2+1} + 3e^{2x}$ *with* $y(0) = 2$.

Exercise 37 *Solve* $\dfrac{dy}{dx} = \dfrac{1}{x^2-1}$ *for* $y(0) = 0$. *(This requires partial fractions or hyperbolic trigonometric functions.)*

Exercise 38 *Solve* $y'' = \sin x$ *for* $y(0) = 0$, $y'(0) = 2$.

Exercise 39 *A spaceship is traveling at the speed* $2t^2 + 1$ km/s *(*t *is time in seconds). It is pointing directly away from earth and at time* $t = 0$ *it is 1000 kilometers from earth. How far from earth is it at one minute from time* $t = 0$*?*

Exercise 40 *Sid is in a car traveling at speed* $10t + 70$ *miles per hour away from Las Vegas, where* t *is in hours. At* $t = 0$*, Sid is 10 miles away from Vegas. How far from Vegas is Sid 2 hours later?*

Exercise 41 *Solve* $\dfrac{dx}{dt} = \sin(t^2) + t$, $x(0) = 20$. *It is OK to leave your answer as a definite integral.*

Exercise 42 *Solve* $\dfrac{dy}{dt} = e^{t^2} + \sin(t)$, $y(0) = 4$. *The answer can be left as a definite integral.*

Author(s): Matthew Charnley and Jason Nowell

Exercise 43 *A dropped ball accelerates downwards at a constant rate 9.8 meters per second squared. Set up the differential equation for the height above ground h in meters. Then supposing $h(0) = 100$ meters, how long does it take for the ball to hit the ground.*

Exercise 44 *The rate of change of the volume of a snowball that is melting is proportional to the surface area of the snowball. Suppose the snowball is perfectly spherical. The volume (in centimeters cubed) of a ball of radius r centimeters is $(4/3)\pi r^3$. The surface area is $4\pi r^2$. Set up the differential equation for how the radius r is changing. Then, suppose that at time $t = 0$ minutes, the radius is 10 centimeters. After 5 minutes, the radius is 8 centimeters. At what time t will the snowball be completely melted?*

Exercise 45 *Find the general solution to $y'''' = 0$. How many distinct constants do you need?*

Slope fields

Stuff about Separable equations

As we said, the general first order equation we are studying looks like

$$y' = f(x, y).$$

A lot of the time, we cannot simply solve these kinds of equations explicitly, because our direct integration method only works when the equation is of the form $y' = f(x)$, which we could integrate directly. In these more complicated cases, it would be nice if we could at least figure out the shape and behavior of the solutions, or find approximate solutions.

Slope fields

Suppose that we have a solution to the equation $y' = f(x, y)$ with $y(x_0) = y_0$. What does the fact that this solves the differential equation tell us about the solution? It tells us that the derivative of the solution at this point will be $f(x_0, y_0)$. Graphically, the derivative gives the slope of the solution, so it means that the solution will pass through the point (x_0, y_0) and will have slope $f(x_0, y_0)$. For example, if $f(x, y) = xy$, then at point $(2, 1.5)$ we draw a short line of slope $xy = 2 \times 1.5 = 3$. So, if $y(x)$ is a solution and $y(2) = 1.5$, then the equation mandates that $y'(2) = 3$. See **Normally a reference to a previous figure goes here.**.

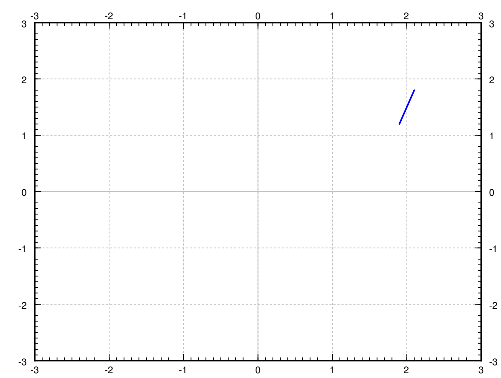


Figure 3: The slope $y' = xy$ at $(2, 1.5)$.

To get an idea of how solutions behave, we draw such lines at lots of points in the plane, not just the point $(2, 1.5)$. We would ideally want to see the slope at every point, but that is just not possible. Usually we pick a grid of points fine enough so that it shows the behavior, but not too fine so that we can still recognize the individual lines. We call this picture the *slope field* of the equation. See **Normally a reference to a previous figure goes here.** for the slope field of the equation $y' = xy$. Usually in practice, one does not do this by hand, but has a computer do the drawing.

The idea of a slope field is that it tells us how the graph of the solution should be sloped, or should curve, if it passed through a given point. Having a wide variety of slopes plotted in our slope field gives an idea of how all of the solutions behave for a bunch of different initial conditions. Which curve we want in particular, and where we should start the curve, depends on the initial condition.

Suppose we are given a specific initial condition $y(x_0) = y_0$. A solution, that is, the graph of the solution, would be a curve that follows the slopes we drew, starting from the point (x_0, y_0). For a few sample solutions, see **Normally a reference to a previous figure goes here.**. It is easy to roughly sketch (or at least imagine) possible solutions in the slope field, just from looking at the slope field itself. You simply sketch a line that roughly fits the little line segments and goes through your initial condition. The graph should "flow" along the little slopes that are on the slope field.

By looking at the slope field we get a lot of information about the behavior of solutions without having to solve the equation. For example, in **Normally a reference to a previous figure goes here.** we see what the solutions do when the initial conditions are $y(0) > 0$, $y(0) = 0$ and $y(0) < 0$. A small change in the initial condition causes quite different behavior. We see this behavior just from the slope field and imagining what solutions ought to do.

We see a different behavior for the equation $y' = -y$. The slope field and a few solutions is in see **Normally a reference to a previous figure goes here.**. If we think of moving from left to right (perhaps x is time and time is usually increasing), then we see that no matter what $y(0)$ is, all solutions tend to zero as x tends to infinity. Again that behavior is clear from simply looking at the slope field itself.

Learning outcomes: Identify or sketch a slope field for a first order differential equation Use the slope field to determine the trajectory of a solution to a differential equation.

Author(s): Matthew Charnley and Jason Nowell

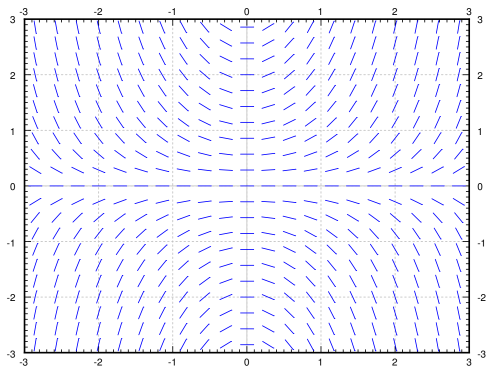


Figure 4: Slope field of $y' = xy$.

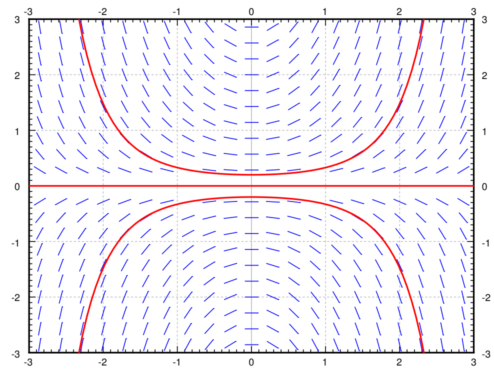


Figure 5: Slope field of $y' = xy$ with a graph of solutions satisfying $y(0) = 0.2$, $y(0) = 0$, and $y(0) = -0.2$.

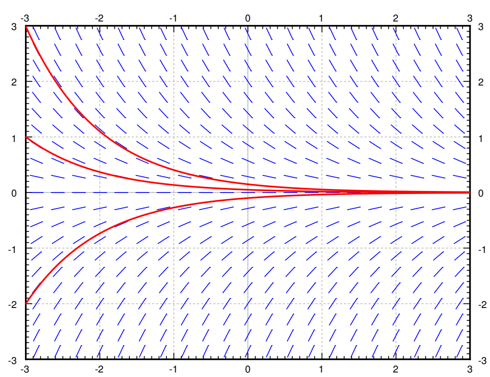


Figure 6: Slope field of $y' = -y$ with a graph of a few solutions.

`firstOrder/practice/slope-fields-practice1.tex`

Practice for Slope Fields

Why?

Exercise 46 *Sketch slope field for* $y' = e^{x-y}$*. How do the solutions behave as* x *grows? Can you guess a particular solution by looking at the slope field?*

Exercise 47 *Sketch the slope field of* $y' = y^3$*. Can you visually find the solution that satisfies* $y(0) = 0$*?*

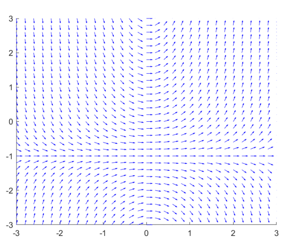
Exercise 48 *Sketch slope field for* $y' = x^2$*.*

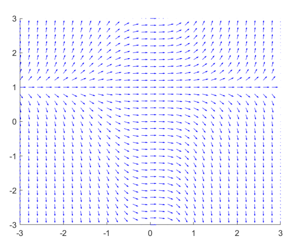
Exercise 49 *Sketch slope field for* $y' = y^2$*.*

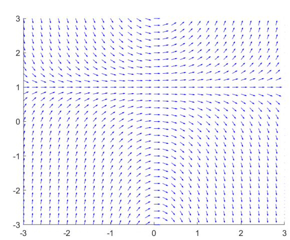
Exercise 50 *For each of the following differential equations, sketch out a slope field on* $-3 < x < 3$ *and* $-3 < y < 3$ *and determine the overall behavior of the solutions to the equation as* $t \to \infty$*. If this fact depends on the value of the solution at* $t = 0$*, explain how it changes.*

a) $\dfrac{dy}{dx} = 3 - 2y$ b) $\dfrac{dy}{dx} = 1 + y$ c) $\dfrac{dy}{dx} = y - 1$ d) $\dfrac{dy}{dx} = -2 - y$

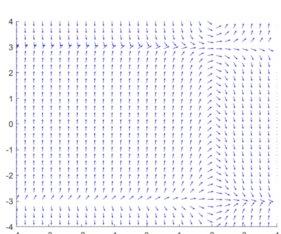
Exercise 51 *Which of the following slope fields corresponds to the differential equation* $\dfrac{dy}{dt} = t(y-1)$*. Explain your reasoning.*

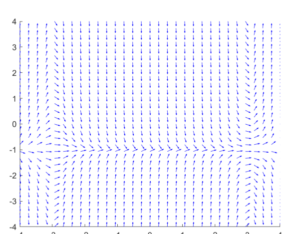
a) 

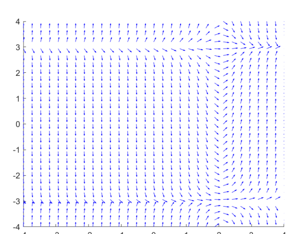
b) 

c) 

Exercise 52 *Which of the following slope fields corresponds to the differential equation* $\dfrac{dy}{dt} = (2-t)(y^2-9)$*. Explain your reasoning.*

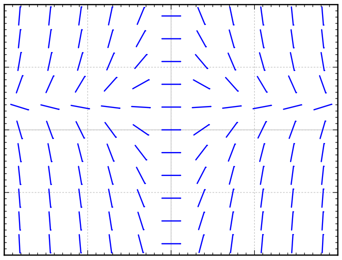
a) 

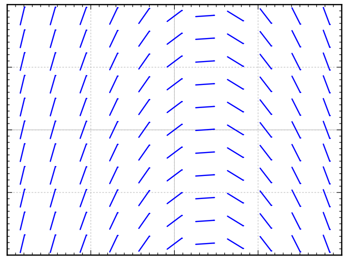
b) 

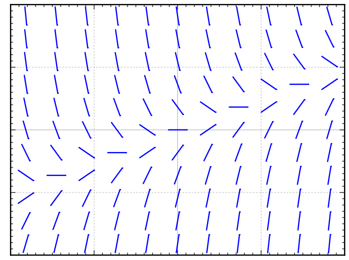
c) 

Author(s): Matthew Charnley and Jason Nowell

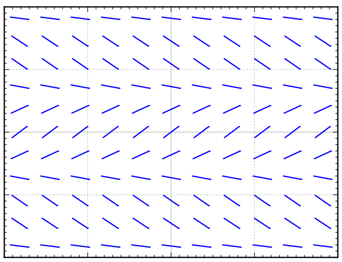
Exercise 53 *Match equations $y' = 1 - x$, $y' = x - 2y$, $y' = x(1 - y)$ to slope fields. Justify.*

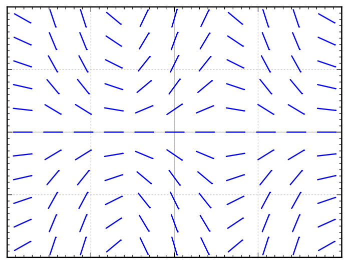
a) 

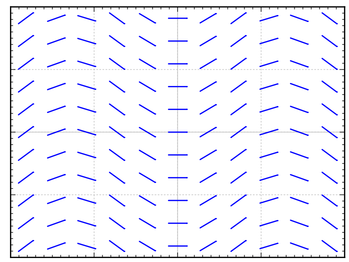
b) 

c) 

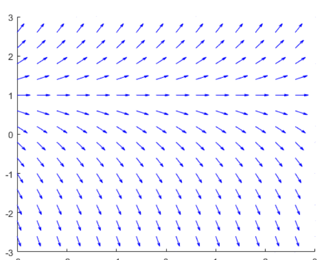
Exercise 54 *Match equations $y' = \sin x$, $y' = \cos y$, $y' = y\cos(x)$ to slope fields. Justify.*

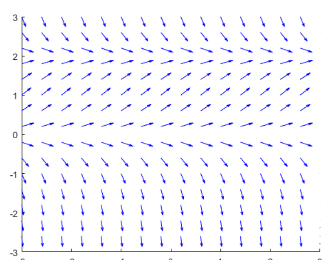
a) 

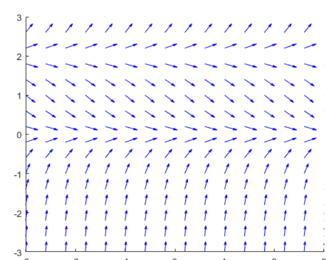
b) 

c) 

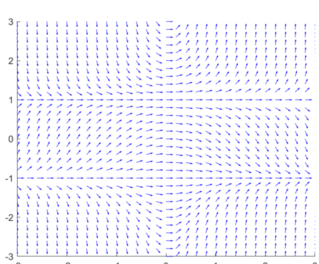
Exercise 55 *Match equations $y' = y(y - 2)$, $y' = y - 1$, $y' = y(2 - y)$ to slope fields. Justify.*

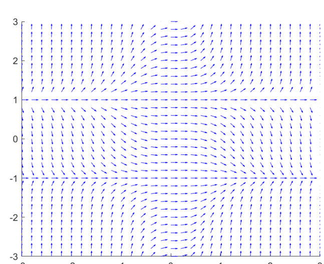
a) 

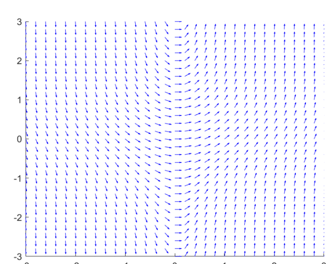
b) 

c) 

Exercise 56 *Match equations $y' = t(y^2 + 1)$, $y' = t(y^2 - 1)$, $y' = t^2(y^2 - 1)$ to slope fields. Justify.*

a) 

b) 

c) 

Exercise 57 *The slope field for the differential equation $y' = (3 - y)(y + 2)$ is below. If we find the solution to this differential equation with initial condition, $y(0) = 1$, what will happen to the solution as $t \to \infty$? Use the slope field and your knowledge of the equation to determine the long-time behavior of this solution.*

Exercise 58 *The slope field for the differential equation $y' = (t-2)(y+4)(y-3)$ is below. If we find the solution to this differential equation with initial condition, $y(0) = 1$, what will happen to the solution as $t \to \infty$? Use the slope field and your knowledge of the equation to determine the long-time behavior of this solution.*

Exercise 59 *The slope field for the differential equation $y' = (y+1)(y+4)$ is below. If we find the solution to this differential equation with initial condition, $y(0) = 1$, what will happen to the solution as $t \to \infty$? Use the slope field and your knowledge of the equation to determine the long-time behavior of this solution.*

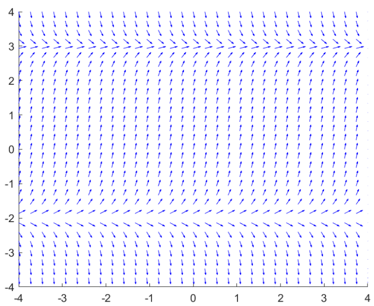


Figure 7: Exercise

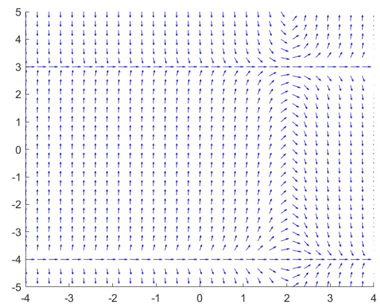


Figure 8: Exercise

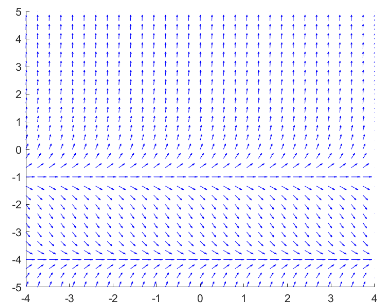


Figure 9: Exercise

Exercise 60 *Take $y' = f(x,y)$, $y(0) = 0$, where $f(x,y) > 1$ for all x and y. If the solution exists for all x, can you say what happens to $y(x)$ as x goes to positive infinity? Explain.*

Exercise 61 *Suppose $y' = f(x,y)$. What will the slope field look like, explain and sketch an example, if you know the following about $f(x,y)$:*

a) *f does not depend on y.*

b) *f does not depend on x.*

c) *$f(t,t) = 0$ for any number t.*

d) *$f(x,0) = 0$ and $f(x,1) = 1$ for all x.*

Exercise 62 *Describe what each of the following facts about the function $f(x,y)$ tells you about the slope field for the differential equation $y' = f(x,y)$.*

a) *$f(2,y) = 0$ for all y*

b) *$f(x,-x) = 0$ for all x*

c) *$f(x,x) = 1$ for all x*

d) *$f(x,-1) = 0$ for all x*

Separable equations

Stuff about Separable equations

As mentioned in § , when a differential equation is of the form $y' = f(x)$, we can just integrate: $y = \int f(x)\,dx + C$. Unfortunately this method no longer works for the general form of the equation $y' = f(x, y)$. Integrating both sides yields

$$y = \int f(x, y)\,dx + C.$$

Notice the dependence on y in the integral. Since y is a function of x, this expression is really of the form

$$y = \int f(x, y(x))\ dx + C$$

and without knowing what $y(x)$ is in advance (which we don't, because that's what we are trying to solve for) we can't compute this integral. Note that while you may have seen integrals of the form

$$\int f(x, y)\ dx$$

in Calculus 3, this is not the same situation. In that class, x and y were both independent variables, so we could integrate this expression in x, treating y as a constant. However, here y is a function of x, so they are not both independent variables and y can not be treated like a constant. If y is a function of x and any y shows up in the integral, you can not compute it.

Separable equations

One particular type of differential equation that we can evaluate using a technique very similar to direct integration is separable equations.

Definition 5. *We say a differential equation is* separable *if we can write it as*

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

for some functions $f(x)$ and $g(y)$.

Let us write the equation in the Leibniz notation

$$\frac{dy}{dx} = f(x)g(y).$$

Then we rewrite the equation as

$$\frac{dy}{g(y)} = f(x)\,dx.$$

It looks like we just separated the derivative as a fraction. The actual reasoning here is the differential from Calculus 1. This is the fact that for y a function of x, we know that

$$dy = \frac{dy}{dx}dx.$$

This means that we can take the equation

$$\frac{dy}{dx} = f(x)g(y),$$

rearrange it as

$$\frac{1}{g(y)}\frac{dy}{dx} = f(x)$$

Learning outcomes: Identify when a differential equation is separable Find the general solution of a separable differential equation Solve initial value problems for separable differential equations.

Author(s): Matthew Charnley and Jason Nowell

and then multiply both sides by dx to get

$$\frac{1}{g(y)}\frac{dy}{dx}\,dx = f(x)dx$$

which leads to the rewritten equation above. Both sides look like something we can integrate. We obtain

$$\int \frac{dy}{g(y)} = \int f(x)\,dx + C.$$

If we can find closed form expressions for these two integrals, we can, perhaps, solve for y.

Example 8. *Solve the equation*

$$y' = xy.$$

Solution: Note that $y = 0$ is a solution. We will remember that fact and assume $y \neq 0$ from now on, so that we can divide by y. Write the equation as $\frac{dy}{dx} = xy$. Then

$$\int \frac{dy}{y} = \int x\,dx + C.$$

We compute the antiderivatives to get

$$\ln|y| = \frac{x^2}{2} + C,$$

or

$$|y| = e^{\frac{x^2}{2}+C} = e^{\frac{x^2}{2}}e^C = De^{\frac{x^2}{2}},$$

where $D > 0$ is some constant. Because $y = 0$ is also a solution and because of the absolute value we can write:

$$y = De^{\frac{x^2}{2}},$$

for any number D (including zero or negative).

We check:

$$y' = Dxe^{\frac{x^2}{2}} = x\left(De^{\frac{x^2}{2}}\right) = xy.$$

Yay!

One particular case in which this method works very well is if the function $f(x, y)$ is only a function of y. With this, we can explicitly complete the solution to equations like

$$y' = ky,$$

reaching the solution $y(x) = e^{kx}$.

We should be a little bit more careful with this method. You may be worried that we integrated in two different variables. We seemingly did a different operation to each side. Let us work through this method more rigorously. Take

$$\frac{dy}{dx} = f(x)g(y).$$

We rewrite the equation as follows. Note that $y = y(x)$ is a function of x and so is $\frac{dy}{dx}$!

$$\frac{1}{g(y)}\frac{dy}{dx} = f(x).$$

We integrate both sides with respect to x:

$$\int \frac{1}{g(y)}\frac{dy}{dx}\,dx = \int f(x)\,dx + C.$$

We use the change of variables formula (substitution) on the left hand side:

$$\int \frac{1}{g(y)}\,dy = \int f(x)\,dx + C.$$

And we are done.

However, in some cases there are some special solutions to these problems as well that don't fit the same formula. Assume we have

$$\frac{dy}{dx} = f(x)g(y)$$

and we have a value y_0 such that $g(y_0) = 0$. Then, the function $y(x) = y_0$ is a solution, provided $f(x)$ is defined everywhere. (Plug this in and check!) This fills in the issue for having $\frac{1}{g(y)}$ in our integral expression, which is not defined when $g(y) = 0$. These are called *singular solutions*, and the next example will showcase one of them.

Implicit solutions

We sometimes get stuck even if we can do the integration. Consider the separable equation

$$y' = \frac{xy}{y^2+1}.$$

We separate variables,

$$\frac{y^2+1}{y}\,dy = \left(y + \frac{1}{y}\right)\,dy = x\,dx.$$

We integrate to get

$$\frac{y^2}{2} + \ln|y| = \frac{x^2}{2} + C,$$

or perhaps the easier looking expression (where $D = 2C$)

$$y^2 + 2\ln|y| = x^2 + D.$$

It is not easy to find the solution explicitly as it is hard to solve for y. We, therefore, leave the solution in this form and call it an *implicit solution*. It is still easy to check that an implicit solution satisfies the differential equation. In this case, we differentiate with respect to x, and remember that y is a function of x, to get

$$y'\left(2y + \frac{2}{y}\right) = 2x.$$

Multiply both sides by y and divide by $2(y^2+1)$ and you will get exactly the differential equation. We leave this computation to the reader.

If you have an implicit solution, and you want to compute values for y, you might have to be tricky. You might get multiple solutions y for each x, so you have to pick one. Sometimes you can graph x as a function of y, and then flip your paper. Sometimes you have to do more.

Computers are also good at some of these tricks. More advanced mathematical software usually has some way of plotting solutions to implicit equations, which makes these solutions just as good for visualizing or graphing as explicit solutions. For example, for $C = 0$ if you plot all the points (x, y) that are solutions to $y^2 + 2\ln|y| = x^2$, you find the two curves in **Normally a reference to a previous figure goes here.**. This is not quite a graph of a function. For each x there are two choices of y. To find a function you would have to pick one of these two curves. You pick the one that satisfies your initial condition if you have one. For example, the top curve satisfies the condition $y(1) = 1$. So for each C we really got two solutions. As you can see, computing values from an implicit solution can be somewhat tricky, but sometimes, an implicit solution is the best we can do.

The equation above also has the solution $y = 0$. Since our function

$$g(y) = \frac{y}{y^2+1}$$

is zero at $y = 0$, and gives an additional solution to the problem. The function $y(x) = 0$ satisfies $y'(x) = 0$ and $\frac{xy}{y^2+1} = 0$ for all x, which is the right-hand side of the equation. So the general solution is

$$y^2 + 2\ln|y| = x^2 + C, \qquad \text{and} \qquad y = 0.$$

These outlying solutions such as $y = 0$ are sometimes called *singular solutions*, as mentioned previously.

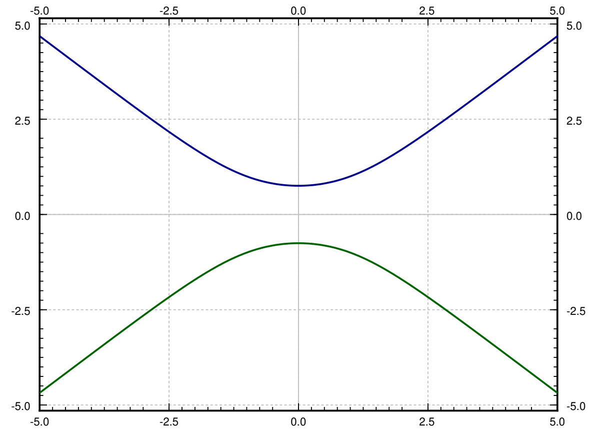


Figure 10: The implicit solution $y^2 + 2\ln|y| = x^2$ to $y' = \dfrac{xy}{y^2+1}$.

Examples of separable equations

Example 9. *Solve* $x^2y' = 1 - x^2 + y^2 - x^2y^2$, $y(1) = 0$.

Solution: Factor the right-hand side

$$x^2y' = (1-x^2)(1+y^2).$$

Separate variables, integrate, and solve for y:

$$\begin{aligned}
\frac{y'}{1+y^2} &= \frac{1-x^2}{x^2}, \\
\frac{y'}{1+y^2} &= \frac{1}{x^2} - 1, \\
\arctan(y) &= \frac{-1}{x} - x + C, \\
y &= \tan\left(\frac{-1}{x} - x + C\right).
\end{aligned}$$

Solve for the initial condition, $0 = \tan(-2 + C)$ to get $C = 2$ (or $C = 2 + \pi$, or $C = 2 + 2\pi$, etc.). The particular solution we seek is, therefore,

$$y = \tan\left(\frac{-1}{x} - x + 2\right).$$

Example 10. *Bob made a cup of coffee, and Bob likes to drink coffee only once reaches 60 degrees Celsius and will not burn him. Initially at time $t = 0$ minutes, Bob measured the temperature and the coffee was 89 degrees Celsius. One minute later, Bob measured the coffee again and it had 85 degrees. The temperature of the room (the ambient temperature) is 22 degrees. When should Bob start drinking?*

Solution: Let T be the temperature of the coffee in degrees Celsius, and let A be the ambient (room) temperature, also in degrees Celsius. Newton's law of cooling states that the rate at which the temperature of the coffee is changing is proportional to the difference between the ambient temperature and the temperature of the coffee. That is,

$$\frac{dT}{dt} = k(A - T),$$

for some constant k. For our setup $A = 22$, $T(0) = 89$, $T(1) = 85$. We separate variables and integrate (let C and D

denote arbitrary constants):

$$\begin{aligned} \frac{1}{T-A}\,\frac{dT}{dt} &= -k, \\ \ln(T-A) &= -kt + C, \qquad \text{(note that } T-A>0\text{)} \\ T-A &= D\,e^{-kt}, \\ T &= A + D\,e^{-kt}. \end{aligned}$$

That is, $T = 22 + D\,e^{-kt}$. We plug in the first condition: $89 = T(0) = 22 + D$, and hence $D = 67$. So $T = 22 + 67\,e^{-kt}$. The second condition says $85 = T(1) = 22 + 67\,e^{-k}$. Solving for k we get $k = -\ln \frac{85-22}{67} \approx 0.0616$. Now we solve for the time t that gives us a temperature of 60 degrees. Namely, we solve

$$60 = 22 + 67e^{-0.0616t}$$

to get $t = -\frac{\ln \frac{60-22}{67}}{0.0616} \approx 9.21$ minutes. So Bob can begin to drink the coffee at just over 9 minutes from the time Bob made it. That is probably about the amount of time it took us to calculate how long it would take. See **Normally a reference to a previous figure goes here.**.

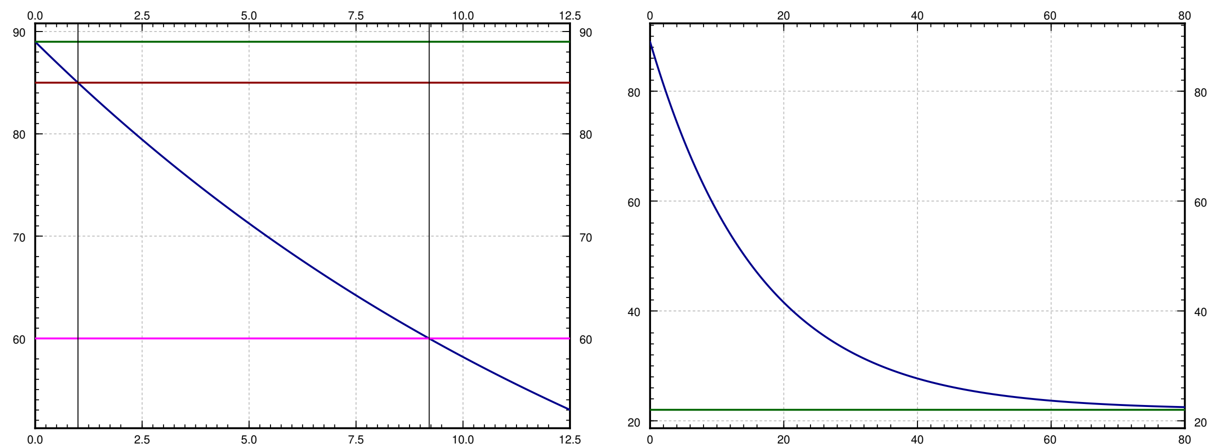


Figure 11: Graphs of the coffee temperature function $T(t)$. On the left, horizontal lines are drawn at temperatures 60, 85, and 89. Vertical lines are drawn at $t = 1$ and $t = 9.21$. Notice that the temperature of the coffee hits 85 at $t = 1$, and 60 at $t \approx 9.21$. On the right, the graph is over a longer period of time, with a horizontal line at the ambient temperature 22.

Example 11. *Find the general solution to* $y' = \frac{-xy^2}{3}$ *(including singular solutions).*

Solution: First note that $y = 0$ is a solution (a singular solution). Now assume that $y \neq 0$.

$$\begin{aligned} \frac{-3}{y^2} y' &= x, \\ \frac{3}{y} &= \frac{x^2}{2} + C, \\ y &= \frac{3}{x^2/2 + C} = \frac{6}{x^2 + 2C}. \end{aligned}$$

So the general solution is,

$$y = \frac{6}{x^2 + 2C}, \qquad \text{and} \qquad y = 0.$$

Example 12. *Find the general solution to*

$$\frac{dy}{dx} = (x^2 + e^x)(y^2 - 3y - 4).$$

Solution: Using the methods of separable equations, we can rewrite this differential equation as

$$\frac{dy}{y^2 - 3y - 4} = (x^2 + e^x)\ dx$$

and we can integrate both sides to solve. This leads to

$$\int \frac{dy}{y^2 - 3y + 4} = \int x^2 + e^x\ dx.$$

The right-hand side of this can be integrated normally to give

$$\int x^2 + e^x\ dx = \frac{x^3}{3} + e^x + C$$

and the left-hand side requires partial fractions in order to integrate correctly. If you are not familiar with this technique of partial fractions, it is reviewed in § **??**.

Using the method of partial fractions, we want to rewrite

$$\frac{1}{y^2 - 3y - 4} = \frac{A}{y - 4} + \frac{B}{y + 1}$$

and solve for A and B, which gives

$$\frac{1}{y^2 - 3y - 4} = \frac{1/5}{y - 4} - \frac{1/5}{y + 1}.$$

Therefore, we can compute the integral

$$\int \frac{dy}{y^2 - 3y - 4} = \int \frac{1/5}{y - 4} - \frac{1/5}{y + 1} dy = \frac{1}{5} \ln(|y - 4|) - \frac{1}{5} \ln(|y + 1|) + C.$$

Therefore, we can write the general solution as

$$\frac{1}{5} \ln\left(\frac{|y - 4|}{|y + 1|}\right) = \frac{x^3}{3} + e^x + C.$$

We could solve this out for y as an explicit function, but that is not necessary for a problem like this.

There are also two singular solutions here at $y = 4$ and $y = -1$. Notice that the implicit solution that we found previously is not defined at either of these values, because they involve taking the natural log of 0, which is not defined.

`firstOrder/practice/separable-practice1.tex`

Practice for Separable ODEs

Why?

Exercise 63 *Solve* $y' = y^3$ *for* $y(0) = 1$.

Exercise 64 *Solve* $x' = \frac{1}{x^2}$, $x(1) = 1$.

Exercise 65 *Solve* $y' = (y-1)(y+1)$ *for* $y(0) = 3$. *(Note: Requires partial fractions)*

Exercise 66 *Solve* $x' = \frac{1}{\cos(x)}$, $x(0) = \frac{\pi}{2}$.

Exercise 67 *Solve* $\frac{dy}{dx} = \frac{1}{y+1}$ *for* $y(0) = 0$.

Exercise 68 *Solve* $y' = {}^{x}\!/_{y}$.

Exercise 69 *Solve* $y' = x^2 y$.

Exercise 70 *Consider the differential equation*

$$\frac{dy}{dx} = \frac{2x}{y}$$

a) *Find the general solution as an implicit function.*

b) *Find the solution to this differential equation as an* explicit *function with* $y(1) = 4$.

c) *Find the solution to this differential equation as an* explicit *function with* $y(0) = -2$.

Exercise 71 *Solve* $y' = y^n$, $y(0) = 1$, *where* n *is a positive integer. Hint: You have to consider different cases.*

Exercise 72 *Solve* $\frac{dx}{dt} = (x^2 - 1)\, t$, *for* $x(0) = 0$. *(Note: Requires partial fractions)*

Exercise 73 *Solve* $\frac{dx}{dt} = x \sin(t)$, *for* $x(0) = 1$.

Author(s): Matthew Charnley and Jason Nowell

Exercise 74 *Solve* $y' = 2xy$.

Exercise 75 *Solve* $y' = ye^{2x}$ *with* $y(0) = 4$.

Exercise 76 *Solve* $\dfrac{dy}{dx} = xy + x + y + 1$. *Hint: Factor the right-hand side.*

Exercise 77 *Solve* $x' = 3xt^2 - 3t^2$, $x(0) = 2$.

Exercise 78 *Find the general solution of* $y' = e^x$, *and then* $y' = e^y$.

Exercise 79 *Solve* $xy' = y + 2x^2y$, *where* $y(1) = 1$.

Exercise 80 *Find an implicit solution for* $x' = \dfrac{1}{3x^2+1}$, $x(0) = 1$.

Exercise 81 *Solve* $\dfrac{dy}{dx} = \dfrac{y^2+1}{x^2+1}$, *for* $y(0) = 1$.

Exercise 82 *Find an implicit solution for* $\dfrac{dy}{dx} = \dfrac{x^2+1}{y^2+1}$, *for* $y(0) = 1$.

Exercise 83 *Find an implicit solution to* $y' = \dfrac{\sin(x)}{\cos(y)}$.

Exercise 84 *Find an implicit solution for* $xy' = \dfrac{x^2+1}{y^2-1}$ *with* $y(3) = 2$.

Exercise 85 *Find an explicit solution for* $y' = xe^{-y}$, $y(0) = 1$.

Exercise 86 *Find an explicit solution to* $xy' = y^2$, $y(1) = 1$.

Exercise 87 *Find an explicit solution for* $xy' = e^{-y}$, *for* $y(1) = 1$.

Exercise 88 *Find an explicit solution for* $y' = y^2(x^4+1)$ *with* $y(1) = 2$.

Exercise 89 *Find an explicit solution for* $y' = \dfrac{\cos(x) + 1}{y}$ *with* $y(0) = 4$.

Exercise 90 *Find an explicit solution for* $y' = ye^{-x^2}$, $y(0) = 1$. *It is alright to leave a definite integral in your answer.*

Exercise 91 *Is the equation* $y' = x + y + 1$ *separable? If so, find the general solution, if not, explain why.*

Exercise 92 *Is the equation* $y' = ty^2 + t$ *separable? If so, find the general solution, if not, explain why.*

Exercise 93 *Is the equation* $y' = xy^2 + 3y^2 - 4x - 12$ *separable? If so, find the general solution, if not, explain why. (Note: Requires partial fractions)*

Exercise 94 *Suppose a cup of coffee is at 100 degrees Celsius at time* $t = 0$, *it is at 70 degrees at* $t = 10$ *minutes, and it is at 50 degrees at* $t = 20$ *minutes. Compute the ambient temperature.*

Exercise 95 *Take* ***Normally a reference to a previous example goes here.*** *with the same numbers: 89 degrees at* $t = 0$, *85 degrees at* $t = 1$, *and ambient temperature of 22 degrees. Suppose these temperatures were measured with precision of* ± 0.5 *degrees. Given this imprecision, the time it takes the coffee to cool to (exactly) 60 degrees is also only known in a certain range. Find this range. Hint: Think about what kind of error makes the cooling time longer and what shorter.*

Exercise 96 *A population* x *of rabbits on an island is modeled by* $x' = x - (1/1000)x^2$, *where the independent variable is time in months. At time* $t = 0$, *there are 40 rabbits on the island.*

a) *Find the solution to the equation with the initial condition.*

b) *How many rabbits are on the island in 1 month, 5 months, 10 months, 15 months (round to the nearest integer).*

Linear equations and the integrating factor

Stuff about Linear equations and the integrating factor

One of the most important types of equations we will learn how to solve are the so-called *linear equations*. In fact, the majority of the course is about linear equations. In this section we focus on the *first order linear equation*.

Definition 6. *A first order equation is* linear *if we can put it into the form:*

$$y' + p(x)y = f(x). \tag{8}$$

The word linear means linear in y and y'; no higher powers nor functions of y or y' appear. The dependence on x can be more complicated.

Solutions of linear equations have nice properties. For example, the solution exists wherever $p(x)$ and $f(x)$ are defined, and has the same regularity (read: it is just as nice). We'll see this in detail in § . But most importantly for us right now, there is a method for solving linear first order equations. In § , we saw that we could easily solve equations of the form

$$\frac{dy}{dx} = f(x)$$

because we could directly integrate both sides of the equation, since the left hand side was the derivative of something (in this case, y) and the right side was only a function of x. We want to do the same here, but the something on the left will not be the derivative of just y.

The trick is to rewrite the left-hand side of (8) as a derivative of a product of y with another function. Let $r(x)$ be this other function, and we can compute, by the product rule, that

$$\frac{d}{dx}\Big[r(x)y\Big] = r(x)y' + r'(x)y.$$

Now, if we multiply (8) by the function $r(x)$ on both sides, we get

$$r(x)y' + p(x)r(x)y = f(x)r(x)$$

and the first term on the left here matches the first term from our product rule derivative. To make the second terms match up as well, we need that

$$r'(x) = p(x)r(x).$$

This equation is separable! We can solve for the $r(x)$ here by separating variables to get that

$$\frac{dr}{r} = p(x)\, dx$$

so that

$$\ln |r| = \int p(x)\, dx$$

or

$$r(x) = e^{\int p(x)\, dx}.$$

With this choice of $r(x)$, we get that

$$r(x)y' + r(x)p(x)y = \frac{d}{dx}\Big[r(x)y\Big],$$

so that if we multiply (8) by $r(x)$, we obtain $r(x)y' + r(x)p(x)y$ on the left-hand side, which we can simplify using our product rule derivative above to obtain

$$\frac{d}{dx}\Big[r(x)y\Big] = r(x)f(x).$$

Learning outcomes: Identify a linear first-order differential equation and write a first-order linear equation in standard form Solve initial value problems for first-order linear differential equations by integrating factors Write solutions to first-order linear initial value problems in integral form if needed.

Author(s): Matthew Charnley and Jason Nowell

Now we integrate both sides. The right-hand side does not depend on y and the left-hand side is written as a derivative of a function. Afterwards, we solve for y. The function $r(x)$ is called the *integrating factor* and the method is called the *integrating factor method.*

This method works for any first order linear equation, no matter what $p(x)$ and $f(x)$ are. In general, we can compute:

$$\begin{aligned} y' + p(x)y &= f(x), \\ e^{\int p(x)\,dx} y' + e^{\int p(x)\,dx} p(x)y &= e^{\int p(x)\,dx} f(x), \\ \frac{d}{dx}\left[e^{\int p(x)\,dx} y \right] &= e^{\int p(x)\,dx} f(x), \\ e^{\int p(x)\,dx} y &= \int e^{\int p(x)\,dx} f(x)\,dx + C, \\ y &= e^{-\int p(x)\,dx} \left(\int e^{\int p(x)\,dx} f(x)\,dx + C \right). \end{aligned}$$

Advice: Do not try to remember the formula itself, that is way too hard. It is easier to remember the process and repeat it.

Of course, to get a closed form formula for y, we need to be able to find a closed form formula for the integrals appearing above.

Example 13. *Solve*

$$y' + 2xy = e^{x-x^2}, \qquad y(0) = -1.$$

Solution: First note that $p(x) = 2x$ and $f(x) = e^{x-x^2}$. The integrating factor is $r(x) = e^{\int p(x)\,dx} = e^{x^2}$. We multiply both sides of the equation by $r(x)$ to get

$$\begin{aligned} e^{x^2} y' + 2xe^{x^2} y &= e^{x-x^2} e^{x^2}, \\ \frac{d}{dx}\left[e^{x^2} y \right] &= e^{x}. \end{aligned}$$

We integrate

$$\begin{aligned} e^{x^2} y &= e^{x} + C, \\ y &= e^{x-x^2} + Ce^{-x^2}. \end{aligned}$$

Next, we solve for the initial condition $-1 = y(0) = 1 + C$, so $C = -2$. The solution is

$$y = e^{x-x^2} - 2e^{-x^2}.$$

Note that we do not care which antiderivative we take when computing $e^{\int p(x)dx}$. You can always add a constant of integration, but those constants will not matter in the end.

Exercise 97 *Try it! Add a constant of integration to the integral in the integrating factor and show that the solution you get in the end is the same as what we got above.*

Since we cannot always evaluate the integrals in closed form, it is useful to know how to write the solution in definite integral form. A definite integral is something that you can plug into a computer or a calculator. Suppose we are given

$$y' + p(x)y = f(x), \qquad y(x_0) = y_0.$$

Look at the solution and write the integrals as definite integrals.

$$y(x) = e^{-\int_{x_0}^{x} p(s)ds} \left(\int_{x_0}^{x} e^{\int_{x_0}^{t} p(s)\,ds} f(t)dt + y_0 \right). \tag{9}$$

You should be careful to properly use dummy variables here. If you now plug such a formula into a computer or a calculator, it will be happy to give you numerical answers.

Exercise 98 *Check that $y(x_0) = y_0$ in formula (9).*

Example 14. *Solve the initial value problem*

$$ty' + 4y = t^2 - 1 \qquad y(1) = 3.$$

Solution: In order to solve this equation, we want to put the equation in standard form, which is

$$y' + \frac{4}{t}y = t - \frac{1}{t}.$$

In this form, the coefficient $p(t)$ of y is $p(t) = \frac{4}{t}$, so that the integrating factor is

$$r(t) = e^{\int p(t)\ dt} = e^{\int \frac{4}{t}\ dt} = e^{4\ln(t)}.$$

Since $4\ln(t) = \ln(t^4)$, we have that $r(t) = t^4$. Multiplying both sides of the equation by t^4 gives

$$t^4y' + 4t^3y = t^5 - t^3$$

where the left hand side is $\frac{d}{dt}(t^4y)$. Therefore, we can integrate both sides of the equation in t to give

$$t^4y = \frac{t^6}{6} - \frac{t^4}{4} + C$$

and we can solve out for y as

$$y(t) = \frac{t^2}{6} - \frac{1}{4} + \frac{C}{t^4}.$$

To solve for C using the initial condition, we plug in $t = 1$ to get that we need

$$3 = \frac{1}{6} - \frac{1}{4} + C \qquad C = \frac{37}{12}.$$

Therefore, the solution to the initial value problem is

$$y(t) = \frac{t^2}{6} - \frac{1}{4} + \frac{37/12}{t^4}.$$

Example 15. *Solve the initial value problem*

$$y' + 2xy = 3 \qquad y(0) = 4.$$

Solution: This equation is already in standard form. Since the coefficient of y is $p(x) = 2x$, we know that the integrating factor is

$$r(x) = e^{\int p(x)\ dx} = e^{x^2}.$$

We can multiply both sides of the equation by this integrating factor to give

$$y'e^{x^2} + 2xe^{x^2}y = 3e^{x^2}$$

and then want to integrate both sides. The left-hand side of the equation is $\frac{d}{dx}[e^{x^2}y]$, so the antiderivative of that side is just ye^{x^2}. For the right-hand side, we would need to compute

$$\int 3e^{x^2}\ dx,$$

which does not have a closed-form expression. Therefore, we need to represent this as a definite integral. Since our initial condition gives the value of y at zero, we want to use zero as the bottom limit of the integral. Therefore, we can write the solution as

$$ye^{x^2} = \int_0^x 3e^{s^2}\ ds + C$$

and so can solve for y as

$$y(x) = e^{-x^2} \int_0^x 3e^{s^2}\, ds + Ce^{-x^2}.$$

Plugging in the initial condition gives that

$$y(0) = 4 = e^{-0} \int_0^0 3e^{s^2}\, ds + Ce^{-0} = C.$$

Therefore, the solution to the initial value problem is

$$y(x) = e^{-x^2} \int_0^x 3e^{s^2}\, ds + 4e^{-x^2}.$$

Exercise 99 *Write the solution of the following problem as a definite integral, but try to simplify as far as you can. You will not be able to find the solution in closed form.*

$$y' + y = e^{x^2 - x}, \qquad y(0) = 10.$$

firstOrder/practice/firstOrderLinear-practice1.tex

Practice for Solving First Order Linear ODEs

Why?

In the exercises, feel free to leave answer as a definite integral if a closed form solution cannot be found. If you can find a closed form solution, you should give that.

Exercise 100 *Solve* $y' + xy = x$.

Exercise 101 *Solve* $y' + 6y = e^x$.

Exercise 102 *Solve* $y' + 4y = x^2 e^{-4x}$.

Exercise 103 *Solve* $y' - 3y = xe^x$.

Exercise 104 *Solve* $y' + 3y = e^{4x} - e^{-2x}$ *with* $y(0) = -3$.

Exercise 105 *Solve* $y' - 2y = x + 4$.

Exercise 106 *Solve* $xy' + 4y = x^2 - \frac{1}{x^2}$.

Exercise 107 *Solve* $xy' - 3y = x - 2$ *with* $y(1) = 3$.

Exercise 108 *Solve* $y' - 4y = \cos(3t)$.

Exercise 109 *Solve* $y' + 3x^2 y = x^2$.

Exercise 110 *Solve* $y' + 3x^2 y = \sin(x)\, e^{-x^3}$, *with* $y(0) = 1$.

Exercise 111 *Solve* $y' + \cos(x)y = \cos(x)$.

Author(s): Matthew Charnley and Jason Nowell

Exercise 112 *Solve the IVP* $4ty' + y = 24\sqrt{t};\ y(10000) = 100$.

Exercise 113 *Solve the IVP* $(t^2+1)y' - 2ty = t^2 + 1;\ y(1) = 0$.

Exercise 114 *Solve* $\dfrac{1}{x^2+1}\,y' + xy = 3$, *with* $y(0) = 0$.

Exercise 115 *Solve* $y' + 2\sin(2x)y = 2\sin(2x)$, $y(\pi/2) = 3$.

Exercise 116 *Consider the initial value problem*

$$5y' - 3y = e^{-2t} \qquad y(0) = a$$

for an undetermined value a. *Solve the problem and determine the dependence on the value of* a. *How does the value of the solution as* $t \to \infty$ *depend on the value of* a*?*

Exercise 117 *Find an expression for the general solution to* $y' + 3y = \sin(t^2)$ *with* $y(0) = 2$. *Simplfy your answer as much as possible.*

Existence and Uniqueness of Solutions

Stuff about Existence and Uniqueness of Solutions

If we take the differential equation

$$y' = f(x, y) \qquad y(x_0) = y_0,$$

there are two main questions we want to answer about this equation.

(a) Does a solution exist to the differential equation?

(b) Is there only one solution to the differential equation?

These are more commonly referred to as (a) existence of the solution and (b) uniqueness of the solution. These are especially crucial for equations that we are using to model a physical situation. For physical situations, the solution definitely exists (because the system does something and continues to exist) and the solution is unique, because a given system will always do the same thing given the same setup. Since we know that physical systems obey these properties, the equations we use to model them should have these properties as well. These properties do not necessarily hold for all differential equations, as shown in the examples below.

Example 16. *Attempt to solve:*

$$y' = \frac{1}{x}, \qquad y(0) = 0.$$

Integrate to find the general solution $y = \ln|x| + C$. The solution does not exist at $x = 0$. See ***Normally a reference to a previous figure goes here.****. The equation may have been written as the seemingly harmless $xy' = 1$.*

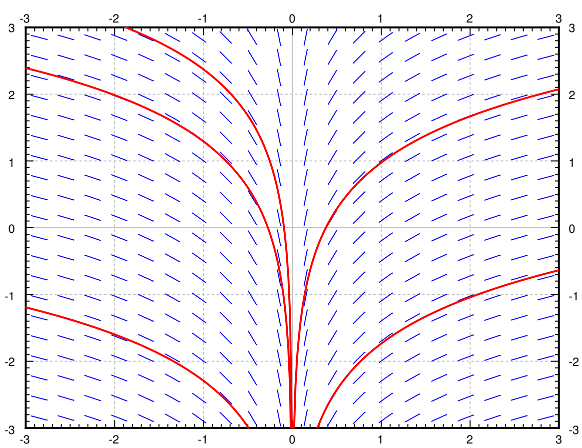


Figure 12: Slope field of $y' = 1/x$.

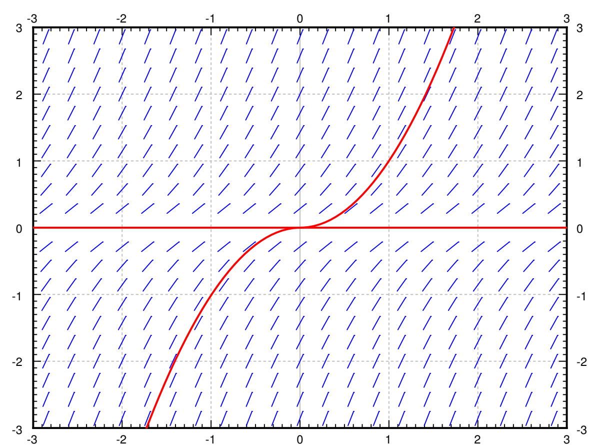


Figure 13: Slope field of $y' = 2\sqrt{|y|}$ with two solutions satisfying $y(0) = 0$.

Example 17. *Solve:*

$$y' = 2\sqrt{|y|}, \qquad y(0) = 0.$$

See ***Normally a reference to a previous figure goes here.****. Note that $y = 0$ is a solution. But another solution is the function*

$$y(x) = \begin{cases} x^2 & \text{if } x \geq 0, \\ -x^2 & \text{if } x < 0. \end{cases}$$

Learning outcomes: Understand the terms existence and uniqueness as they apply to differential equations Find the maximum guaranteed interval of existence for the solution to an initial value problem.

Author(s): Matthew Charnley and Jason Nowell

What we see here is a significant problem for trying to represent physical situations. In the first there is no solution at $x = 0$, so if our physical scenario had wanted one, that would be an issue. Similarly, for the second, we do have solutions, but we have two of them, so we can't use this to predict what is going to happen to a physical situation modeled by this equation over time. So, we need both existence and uniqueness to hold for our modeling equation in order to use differential equations to accurately model situations. Thankfully, these properties do apply to most equations, and we have fairly straight-forward criteria that can be used to determine if these properties are true for a given differential equation. For a first-order linear differential equation, the theorem is fairly straight-forward.

Theorem 1 (exist:lin1thm)**.** *Existence and Uniqueness Theorem Assume that we have the first-order linear differential equation given by*

$$y' + p(x)y = g(x).$$

If $p(x)$ and $g(x)$ are continuous functions on an interval I that contains a point x_0, then for any y-value y_0, the initial value problem

$$y' + p(x)y = g(x) \qquad y(x_0) = y_0$$

has a unique solution. This solution exists and is unique on the entire interval I.

The idea and proof of this theorem comes from the fact that we have an explicit method for solving these equations no matter what p and g are. We can always find an integrating factor for the equation, convert the left-hand side into a product rule term, take a definite integral of both sides, and then solve for y. Since we have this explicit formula, the solution will exist and be defined on the entire interval where the functions p and g are continuous. This also means that we can answer questions about where and for what values of x the solution to a differential equation exists.

Example 18. *Consider the differential equation*

$$(x-1)y' + \frac{1}{x-5}y = e^x$$

What do the existence and uniqueness theorems say about the solution to this differential equation with the initial condition $y(2) = 6$? What about the solution with initial condition $y(-3) = 1$?

Solution: To apply the existence and uniqueness theorem, we need to get the y' term by itself. This results in the differential equation

$$y' + \frac{1}{(x-1)(x-5)}y = \frac{e^x}{x-1}.$$

In order to figure out where this solution exists and is unique, we need to determine where the coefficient functions $p(x)$ and $g(x)$ are continuous. The only two points that we have discontinuities are at $x = 1$ and $x = 5$. Therefore, if we have the initial condition $y(2) = 6$, we start at the x value of 2. Because this equation is linear, it will exist everywhere that these two functions are both continuous containing the point $x = 2$, and since the only discontinuities are at 1 and 5, we know that they are both continuous on $(1, 5)$. This means that we can take $(1, 5)$ as the interval I in the theorem, and know that this solution will exist and be unique on the interval $(1, 5)$.

For the other initial condition, $y(-3) = 1$, we now want an interval where these functions are continuous that contains -3. Again, we only have to avoid $x = 1$ and $x = 5$, so we can take the interval $(-\infty, 1)$ as the interval I in the theorem, and so we know the solution with this initial condition will exist and be unique on $(-\infty, 1)$.

A convenient way to represent this situation is with a number line like that presented in **Normally a reference to a previous figure goes here.**. On this number line, we mark the places where the functions $p(x)$ or $g(x)$ are discontinuous.

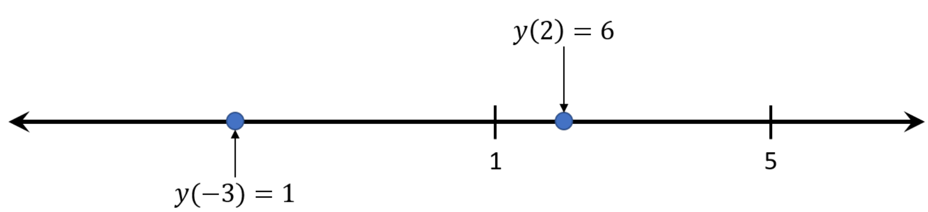


Figure 14: Number line representation of the existence intervals for a differential equation.

To interpret this image, we can mark the initial point on the number line, where the point that we mark is the x coordinate of the initial condition. All of the intervals are in terms of x. Then, the existence and uniqueness theorem says that the solution will exist on the entire interval between any marked points on this number line. From that, we can see that the interval of existence for the initial condition $y(2) = 6$ is $(1, 5)$, and the interval for $y(-3) = 1$ is $(-\infty, 1)$.

For non-linear equations, we don't have an explicit method of getting a solution that works for all equations. This means that we can't fall back on this formula to guarantee existence or uniqueness of solutions. For this reason, we expect to get a result that is not as strong for non-linear equations. Thankfully, we do still get a result, which is known as Picard's theorem[1].

Theorem 2 (slope:picardthm)**.** *Picard's theorem on existence and uniquenessExistence and uniqueness, Picard's theorem: If $f(x, y)$ is continuous (as a function of two variables) and $\frac{\partial f}{\partial y}$ exists and is continuous near some (x_0, y_0), then a solution to*

$$y' = f(x, y), \qquad y(x_0) = y_0,$$

exists (at least for some small interval of x's) and is unique.

The main fact that is "not as strong" about this result is the interval that we get from the theorem. For the linear theorem, we got existence and uniqueness on the entire interval I where p and g are continuous. For the non-linear theorem, we only get existence on *some* interval around the point x_0. Even if $f(x, y)$ and $\frac{\partial f}{\partial y}$ are really nice functions that are continuous everywhere, we can still only guarantee existence on a small interval (that can depend on the initial condition) around the point x_0.

Example 19. *For some constant A, solve:*

$$y' = y^2, \qquad y(0) = A.$$

Solution: We know how to solve this equation. First assume that $A \neq 0$, so y is not equal to zero at least for some x near 0. So $x' = {}^1/_{y^2}$, so $x = {}^{-1}/_{y} + C$, so $y = \frac{1}{C - x}$. If $y(0) = A$, then $C = {}^1/_A$ so

$$y = \frac{1}{{}^1/_A - x}.$$

If $A = 0$, then $y = 0$ is a solution.

For example, when $A = 1$ the solution is

$$y = \frac{1}{1 - x}$$

which goes to infinity, and so "blows up", at $x = 1$. This solution here exists only on the interval $(-\infty, 1)$, and hence, the solution does not exist for all x even if the equation is nice everywhere. The equation $y' = y^2$ certainly looks nice.

However, this fact does not contradict our existence and uniqueness theorem for non-linear equations. The theorem only guarantees that the solution to

$$y' = y^2$$

exists and is unique on *some* interval containing 0. It does not guarantee that the solution exists everywhere that y^2 and its derivative are continuous, only that at each point where this happens, the solution will exist for some interval around that point. The interval $(-\infty, 1)$ is "some interval containing 0", so the theorem still applies and holds here. See the exercises for more detail on how this process works and how we can illustrate the fact that the interval of existence ise "some interval containing 0".

The other main conclusion that we can draw from these theorems is the fact that two different solution curves to a first-order differential equation can not cross, provided the existence and uniqueness theorems hold. If y_1 and y_2 are two different solutions to $y' = f(x, y)$ and the solution curves for $y_1(x)$ and $y_2(x)$ cross, then this means that for some particular value of x_0 and y_0, we have that

$$y_1(x_0) = y_0 \qquad y_2(x_0) = y_0.$$

If we pick x_0 as a starting point, then the fact that the existence and uniqueness theorems hold imply that, at least for some interval around x_0, there is exactly one solution to

$$y' = f(x, y) \qquad y(x_0) = y_0.$$

[1] Named after the French mathematician Charles Émile Picard (1856–1941)

However, both y_1 and y_2 satisfy these two properties. Therefore, y_1 and y_2 must be the same, which doesn't make sense because we assumed they were different. So it is impossible for two different solution curves to cross, provided the existence and uniqueness theorem holds. For a comparison, refer back to **Normally a reference to a previous example goes here.** earlier to see what non-uniqueness looks like, where we do have two solution curves that cross at the point $(0, 0)$.

This fact is useful for analyzing differential equations in general, but will be particularly useful in § in dealing with autonomous equations, where we can use simple solutions to provide boundaries over which other solutions can not cross. This fact will come up again in Chapters **??** and **??** in sketching trajectories for these solutions as well.

Example 20. *Consider the differential equation*

$$\frac{dy}{dt} = (y-3)^2(y+4).$$

(a) *Verify that $y = 3$ is a solution to this differential equation.*

(b) *Assume that we solve this problem with initial condition $y(0) = 1$. Is it possible for this solution to ever reach $y = 4$? Why or why not?*

Solution:

(a) If we take the function $y(t) = 3$, then $y' = 0$, and plugging this into the right hand side also gives 0. Therefore, this function solves the differential equation.

(b) If the solutions starts with $y(0) = 1$, this means that it starts below the line $y = 3$. In order to get up to $y = 4$, the solution would need to cross over the line $y = 3$, which would mean that we have solution curves that cross. However, the function $f(t, y) = (y-3)^2(y+4)$ is continuous everywhere, as is the first derivative $\frac{\partial f}{\partial y} = 2(y-3)(y+4)+(y-3)^2$. Therefore, the existence and uniqueness theorem applies everywhere, and so solution curves can not cross. So, it is not possible for the solution to reach $y = 4$, because this would force solution curves to cross, which we know can not happen.

firstOrder/practice/EUThm-practice1.tex

Practice for Using the EU-Thm

Why?

Exercise 118 *Is it possible to solve the equation* $y' = \dfrac{xy}{\cos x}$ *for* $y(0) = 1$*? Justify.*

Exercise 119 *Is it possible to solve the equation* $y' = y\sqrt{|x|}$ *for* $y(0) = 0$*? Is the solution unique? Justify.*

Exercise 120 *Consider the differential equation* $y' + \dfrac{1}{t-2}y = \dfrac{1}{t+3}$.

a) *Is this equation linear or non-linear?*

b) *What is the maximum guaranteed interval of existence for the solution to this equation with initial condition* $y(0) = 3$*?*

c) *What if we start with the initial condition* $y(4) = 0$*?*

Exercise 121 *Consider the differential equation* $y' + \dfrac{1}{t+2}y = \dfrac{\ln(|t|)}{t-4}$.

a) *Is this equation linear or non-linear?*

b) *What is the maximum guaranteed interval of existence for the solution to this equation with initial condition* $y(-3) = 1$*?*

c) *What if we start with the initial condition* $y(2) = 5$*?*

d) *What happens if we want to start with* $y(4) = 3$*?*

Exercise 122 *Consider the differential equation* $(t+3)y' + t^2 y = \dfrac{1}{t-2}$.

a) *Is this equation linear or non-linear?*

b) *What is the maximum guaranteed interval of existence for the solution to this equation with initial condition* $y(-2) = 1$*?*

c) *What if we start with the initial condition* $y(-4) = 5$*?*

d) *What happens if we want to start with* $y(4) = 2$*?*

Exercise 123 *Consider the differential equation* $y' = y^2$.

Author(s): Matthew Charnley and Jason Nowell

a) *Is this equation linear or non-linear?*

b) *What is the most we can say about the interval of existence for the solution to this equation with initial condition* $y(0) = 1$*?*

c) *Find the solution to this differential equation with* $y(0) = 1$*. Over what values of* x *does this solution exist?*

d) *Find the solution to this differential equation with* $y(0) = 4$*. Over what values of* x *does this solution exist?*

e) *Find the solution to this differential equation with* $y(0) = -2$*. Over what values of* x *does this solution exist?*

f) *Do any of these contradict your answer in (b)?*

Exercise 124 *Consider the differential equation* $y' = y^2 + 4$*.*

a) *Is this equation linear or non-linear?*

b) *What is the most we can say about the interval of existence for the solution to this equation with initial condition* $y(0) = 0$*?*

c) *Find the solution to this differential equation with* $y(0) = 0$*. Over what values of* x *does this solution exist?*

Exercise 125 *Consider the differential equation* $y' = x(y+1)^2$*.*

a) *Is this equation linear or non-linear?*

b) *If we set* $f(x, y) = x(y+1)^2$*, for what values of* x *and* y *are* f *and* $\dfrac{\partial f}{\partial y}$ *continuous?*

c) *What is the most we can say about the interval of existence for the solution to this equation with initial condition* $y(0) = 1$*?*

d) *Find the solution to this differential equation with* $y(0) = 1$*. Over what values of* x *does this solution exist?*

Exercise 126 *Take* $(y - x)y' = 0$*,* $y(0) = 0$*.*

a) *Find two distinct solutions.*

b) *Explain why this does not violate Picard's theorem.*

Exercise 127 *Find a solution to* $y' = |y|$*,* $y(0) = 0$*. Does Picard's theorem apply?*

Exercise 128 *Consider the IVP* $y' \cos t + y \sin t = 1$*;* $y(\pi/6) = 1$*.*

a) *The Existence and Uniqueness Theorem guarantees a unique solution to this IVP on what interval?*

b) *Find this solution explicitly.*

Exercise 129 *Take an equation $y' = (y - 2x)g(x, y) + 2$ for some function $g(x, y)$. Can you solve the problem for the initial condition $y(0) = 0$, and if so what is the solution?*

Exercise 130 *Consider the differential equation $y' = e^x(2 - y)$.*

a) *Verify that $y = 2$ is a solution to this differential equation.*

b) *Assume that we look for the solution with $y(0) = 0$. Is it possible that $y(x) = 3$ for some later time x? Why or why not?*

c) *Based on this, what do we know about the solution with $y(0) = 5$?*

Exercise 131 *Suppose $y' = f(x, y)$ is such that $f(x, 1) = 0$ for every x, f is continuous and $\dfrac{\partial f}{\partial y}$ exists and is continuous for every x and y.*

a) *Guess a solution given the initial condition $y(0) = 1$.*

b) *Can graphs of two solutions of the equation for different initial conditions ever intersect?*

c) *Given $y(0) = 0$, what can you say about the solution. In particular, can $y(x) > 1$ for any x? Can $y(x) = 1$ for any x? Why or why not?*

Exercise 132 *Consider the differential equation $y' = y^2 - 4$.*

a) *Verify that $y = 2$ and $y = -2$ are both solutions to this differential equation.*

b) *Verify that the hypotheses of Picard's theorem are satisfies for this equation.*

c) *Assume that we solve this differential equation with $y(0) = 1$. Is it possible for the solution to reach $y = 3$ at any point? Why or why not?*

d) *Assume that we solve this differential equation with $y(0) = -1$. Is it possible for the solution to reach $y = -4$ at any point? Why or why not?*

Exercise 133 *Is it possible to solve $y' = xy$ for $y(0) = 0$? Is the solution unique?*

Exercise 134 *Is it possible to solve $y' = \dfrac{x}{x^2 - 1}$ for $y(1) = 0$?*

Exercise 135 *Suppose*

$$f(y) = \begin{cases} 0 & \text{if } y > 0, \\ 1 & \text{if } y \le 0. \end{cases}$$

Does $y' = f(y)$, $y(0) = 0$ have a continuously differentiable solution? Does Picard apply? Why, or why not?

Exercise 136 *Consider an equation of the form $y' = f(x)$ for some continuous function f, and an initial condition $y(x_0) = y_0$. Does a solution exist for all x? Why or why not?*

Numerical methods: Euler's method

Stuff about Numerical methods: Euler's method

Unless $f(x, y)$ is of a special form, it is generally very hard if not impossible to get a nice formula for the solution of the problem

$$y' = f(x, y), \qquad y(x_0) = y_0.$$

If the equation can be solved in closed form, we should do that. But what if we have an equation that cannot be solved in closed form? What if we want to find the value of the solution at some particular x? Or perhaps we want to produce a graph of the solution to inspect the behavior. In this section we will learn about the basics of numerical approximation of solutions.

The simplest method for approximating a solution is *Euler's method*[1]. It works as follows: Take x_0 and compute the slope $k = f(x_0, y_0)$. The slope is the change in y per unit change in x. Follow the line for an interval of length h on the x-axis. Hence if $y = y_0$ at x_0, then we say that y_1 (the approximate value of y at $x_1 = x_0 + h$) is $y_1 = y_0 + hk$. Rinse, repeat! Let $k = f(x_1, y_1)$, and then compute $x_2 = x_1 + h$, and $y_2 = y_1 + hk$. Now compute x_3 and y_3 using x_2 and y_2, etc. Consider the equation $y' = y^2/3$, $y(0) = 1$, and $h = 1$. Then $x_0 = 0$ and $y_0 = 1$. We compute

$$\begin{aligned} x_1 &= x_0 + h = 0 + 1 = 1, & y_1 &= y_0 + h\, f(x_0, y_0) = 1 + 1 \cdot 1^2/3 = 4/3 \approx 1.333, \\ x_2 &= x_1 + h = 1 + 1 = 2, & y_2 &= y_1 + h\, f(x_1, y_1) = 4/3 + 1 \cdot \frac{(4/3)^2}{3} = 52/27 \approx 1.926. \end{aligned}$$

We then draw an approximate graph of the solution by connecting the points (x_0, y_0), (x_1, y_1), (x_2, y_2),.... For the first two steps of the method see **Normally a reference to a previous figure goes here.**.

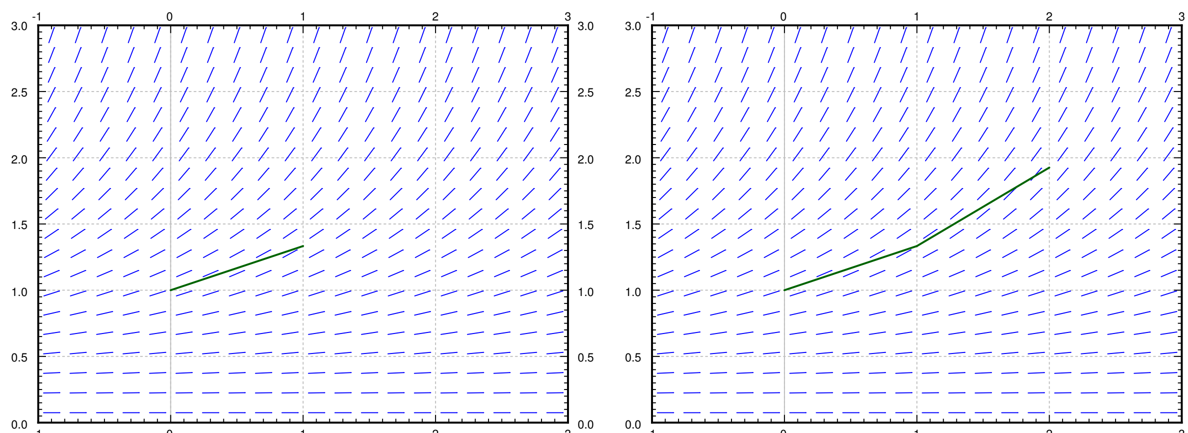


Figure 15: First two steps of Euler's method with $h = 1$ for the equation $y' = \dfrac{y^2}{3}$ with initial conditions $y(0) = 1$.

More abstractly, for any $i = 0, 1, 2, 3, \ldots$, we compute

$$x_{i+1} = x_i + h, \qquad y_{i+1} = y_i + h\, f(x_i, y_i).$$

This can be worked out by hand for a few steps, but the formulas here lend themselves very well to being coded into a looping structure for more involved processes. The line segments we get are an approximate graph of the solution. Generally it is not exactly the solution. See **Normally a reference to a previous figure goes here.** for the plot of the real solution and the approximation.

Learning outcomes: Use Euler's method to numerically approximate solutions to first order differential equations Compute the error in a numerical method using the true solution Compare a variety of numerical methods, including built-in Matlab methods.

Author(s): Matthew Charnley and Jason Nowell

[1]Named after the Swiss mathematician Leonhard Paul Euler (1707–1783). The correct pronunciation of the name sounds more like "oiler."

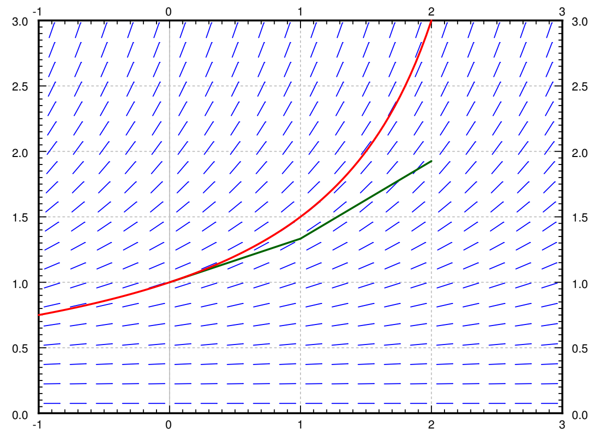


Figure 16: Two steps of Euler's method (step size 1) and the exact solution for the equation $y' = \frac{y^2}{3}$ with initial conditions $y(0) = 1$.

We continue with the equation $y' = {}^{y^2}/_3$, $y(0) = 1$. Let us try to approximate $y(2)$ using Euler's method. In Figures 15 and 16 we have graphically approximated $y(2)$ with step size 1. With step size 1, we have $y(2) \approx 1.926$. The real answer is 3. We are approximately 1.074 off. Let us halve the step size. Computing y_4 with $h = 0.5$, we find that $y(2) \approx 2.209$, so an error of about 0.791. **Normally a reference to a previous table goes here.** gives the values computed for various parameters.

Exercise 137 *Solve this equation exactly and show that $y(2) = 3$.*

The difference between the actual solution and the approximate solution is called the error. We usually talk about just the size of the error and we do not care much about its sign. The point is, we usually do not know the real solution, so we only have a vague understanding of the error. If we knew the error exactly ... what is the point of doing the approximation?

h	Approximate $y(2)$	Error	$\frac{\text{Error}}{\text{Previous error}}$
1	1.92593	1.07407	
0.5	2.20861	0.79139	0.73681
0.25	2.47250	0.52751	0.66656
0.125	2.68034	0.31966	0.60599
0.0625	2.82040	0.17960	0.56184
0.03125	2.90412	0.09588	0.53385
0.015625	2.95035	0.04965	0.51779
0.0078125	2.97472	0.02528	0.50913

Table 1: Euler's method approximation of $y(2)$ where of $y' = {}^{y^2}/_3$, $y(0) = 1$.

Notice that except for the first few times, every time we halved the step size the error approximately halved. This halving of the error is a general feature of Euler's method as it is a *first order method*. There exists an improved Euler method, see the exercises, which is a second order method. A second order method reduces the error to approximately one quarter every time we halve the interval. The meaning of "second" order is the squaring in ${}^1/_4 = {}^1/_2 \times {}^1/_2 = ({}^1/_2)^2$.

To get the error to be within 0.1 of the answer we had to already do 64 steps. To get it to within 0.01 we would have to halve another three or four times, meaning doing 512 to 1024 steps. That is quite a bit to do by hand. The improved Euler method from the exercises should quarter the error every time we halve the interval, so we would have to approximately do half as many "halvings" to get the same error. This reduction can be a big deal. With 10 halvings (starting at $h = 1$) we have 1024 steps, whereas with 5 halvings we only have to do 32 steps, assuming that the error was comparable to start with. A computer may not care about this difference for a problem this simple, but suppose each step would take a second to compute (the function may be substantially more difficult to compute than $y^2/3$). Then the difference is 32 seconds versus about 17 minutes. We are not being altogether fair, a second order method would probably double the time to do each step. Even so, it is 1 minute versus 17 minutes. Next, suppose that we have to repeat such a calculation for different parameters a thousand times. You get the idea.

Note that in practice we do not know how large the error is! How do we know what is the right step size? Well, essentially we keep halving the interval, and if we are lucky, we can estimate the error from a few of these calculations and the assumption that the error goes down by a factor of one half each time (if we are using standard Euler).

Exercise 138 *In the table above, suppose you do not know the error. Take the approximate values of the function in the last two lines, assume that the error goes down by a factor of 2. Can you estimate the error in the last time from this? Does it (approximately) agree with the table? Now do it for the first two rows. Does this agree with the table?*

Let us talk a little bit more about the example $y' = \frac{y^2}{3}$, $y(0) = 1$. Suppose that instead of the value $y(2)$ we wish to find $y(3)$. The results of this effort are listed in **Normally a reference to a previous table goes here.** for successive halvings of h. What is going on here? Well, you should solve the equation exactly and you will notice that the solution does not exist at $x = 3$. In fact, the solution goes to infinity when you approach $x = 3$.

h	Approximate $y(3)$
1	3.16232
0.5	4.54329
0.25	6.86079
0.125	10.80321
0.0625	17.59893
0.03125	29.46004
0.015625	50.40121
0.0078125	87.75769

Table 2: Attempts to use Euler's to approximate $y(3)$ where of $y' = y^2/3$, $y(0) = 1$.

Another case where things go bad is if the solution oscillates wildly near some point. The solution may exist at all points, but even a much better numerical method than Euler would need an insanely small step size to approximate the solution with reasonable precision. And computers might not be able to easily handle such a small step size.

In real applications we would not use a simple method such as Euler's. The simplest method that would probably be used in a real application is the standard Runge–Kutta method (see exercises). That is a fourth order method, meaning that if we halve the interval, the error generally goes down by a factor of 16 (it is fourth order as $1/16 = 1/2 \times 1/2 \times 1/2 \times 1/2$).

Choosing the right method to use and the right step size can be very tricky. There are several competing factors to consider.

- Computational time: Each step takes computer time. Even if the function f is simple to compute, we do it many times over. Large step size means faster computation, but perhaps not the right precision.
- Roundoff errors: Computers only compute with a certain number of significant digits. Errors introduced by rounding numbers off during our computations become noticeable when the step size becomes too small relative to the quantities we are working with. So reducing step size may in fact make errors worse. There is a certain optimum step size such that the precision increases as we approach it, but then starts getting worse as we make our step size smaller still. Trouble is: this optimum may be hard to find.
- Stability: Certain equations may be numerically unstable. What may happen is that the numbers never seem to stabilize no matter how many times we halve the interval. We may need a ridiculously small interval size, which may not be practical due to roundoff errors or computational time considerations. Such problems are sometimes called *stiff* stiff problem. In the worst case, the numerical computations might be giving us bogus numbers that look like a

correct answer. Just because the numbers seem to have stabilized after successive halving, does not mean that we must have the right answer.

We have seen just the beginnings of the challenges that appear in real applications. Numerical approximation of solutions to differential equations is an active research area for engineers and mathematicians. For example, the general purpose method used for the ODE solver in Matlab and Octave (as of this writing) is a method that appeared in the literature only in the 1980s.

The method used in Matlab and Octave is a fair bit different from the methods discussed previously. We don't need to go too much in detail about it, but some information will be helpful. The main difference that will be visible when running these methods is that they are *adaptive* method. This means that they adjust the step-size based on what the differential equation looks like at a given point. Euler's method, along with the improved Euler and Runge-Kutta methods, is a fixed step-size method, where the steps are always the same no matter what. Adaptive methods are harder to write and optimize, but can solve many problems faster because the adaptive nature of the method allows them to get similar accuracy to fixed step methods, but at many fewer steps. In the example below, the initial value problem

$$\frac{dy}{dt} = y \qquad y(0) = 1$$

is solved with an Euler's method and Matlab's built-in `ode45` method. Both of the solutions are plotted along with the actual solution $y = e^t$

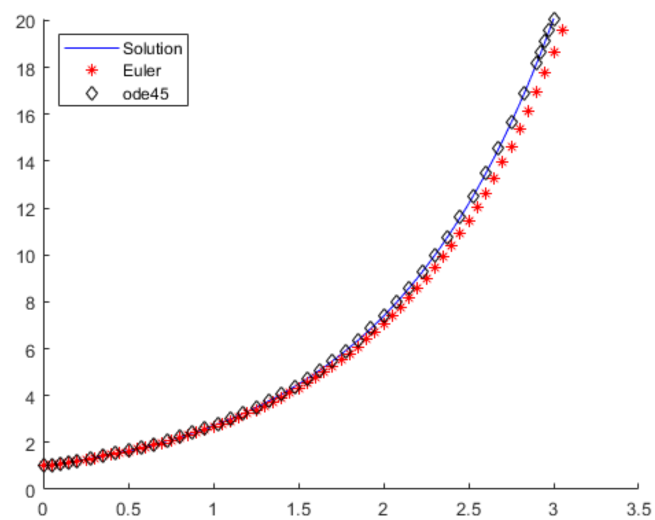


Figure 17: Comparison of the solution from Euler's Method and ode45 to the actual solution of $\frac{dy}{dt} = y$.

The Euler's method takes 60 steps in this computation, but is still not as accurate as the `ode45` method, which only takes 45 steps. In addition, the black diamonds, representing the different values computed by `ode45` are not evenly spaced, illustrating the adaptive nature of this solver, while the red stars are all evenly spaced in the t-direction, as is expected from Euler's method.

firstOrder/practice/numerical-Euler-practice1.tex

Practice for Euler Method

Why?

Exercise 139 *Consider* $\frac{dx}{dt} = (2t - x)^2$, $x(0) = 2$. *Use Euler's method with step size* $h = 0.5$ *to approximate* $x(1)$.

Exercise 140 *Consider the differential equation* $\frac{dy}{dt} = t^2 - 3y + 1$ *with* $y(1) = 4$. *Approximate the solution at* $t = 3$ *using Euler's method with a step size of* $h = 1$ *and* $h = 0.5$. *Compare these values with the actual solution at* $t = 3$.

Exercise 141 *Consider the differential equation* $\frac{dy}{dt} = 2ty + y^2$ *with* $y(0) = 1$. *Approximate the solution at* $t = 2$ *using Euler's method with a step size of* $h = 1$ *and* $h = 0.5$.

Exercise 142 *Consider* $\frac{dx}{dt} = t - x$, $x(0) = 1$.

a) *Use Euler's method with step sizes* $h = 1, {}^1/_2, {}^1/_4, {}^1/_8$ *to approximate* $x(1)$.
b) *Solve the equation exactly.*
c) *Describe what happens to the errors for each* h *you used. That is, find the factor by which the error changed each time you halved the interval.*

Exercise 143 *Let* $x' = \sin(xt)$, *and* $x(0) = 1$. *Approximate* $x(1)$ *using Euler's method with step sizes 1, 0.5, 0.25. Use a calculator and compute up to 4 decimal digits.*

Exercise 144 *Approximate the value of* e *by looking at the initial value problem* $y' = y$ *with* $y(0) = 1$ *and approximating* $y(1)$ *using Euler's method with a step size of* 0.2.

Exercise 145 *Let* $x' = 2t$, *and* $x(0) = 0$.

a) *Approximate* $x(4)$ *using Euler's method with step sizes 4, 2, and 1.*
b) *Solve exactly, and compute the errors.*
c) *Compute the factor by which the errors changed.*

Exercise 146 *Let* $x' = xe^{xt+1}$, *and* $x(0) = 0$.

a) *Approximate* $x(4)$ *using Euler's method with step sizes 4, 2, and 1.*
b) *Guess an exact solution based on part a) and compute the errors.*

Author(s): Matthew Charnley and Jason Nowell

Exercise 147 *Example of numerical instability: Take $y' = -5y$, $y(0) = 1$. We know that the solution should decay to zero as x grows. Using Euler's method, start with $h = 1$ and compute y_1, y_2, y_3, y_4 to try to approximate $y(4)$. What happened? Now halve the interval. Keep halving the interval and approximating $y(4)$ until the numbers you are getting start to stabilize (that is, until they start going towards zero). Note: You might want to use a calculator.*

There is a simple way to improve Euler's method to make it a second order method by doing just one extra step. Consider $\frac{dy}{dx} = f(x, y)$, $y(x_0) = y_0$, and a step size h. What we do is to pretend we compute the next step as in Euler, that is, we start with (x_i, y_i), we compute a slope $k_1 = f(x_i, y_i)$, and then look at the point $(x_i + h, y_i + k_1 h)$. Instead of letting our new point be $(x_i + h, y_i + k_1 h)$, we compute the slope at that point, call it k_2, and then take the average of k_1 and k_2, hoping that the average is going to be closer to the actual slope on the interval from x_i to $x_i + h$. And we are correct, if we halve the step, the error should go down by a factor of $2^2 = 4$. To summarize, the setup is the same as for regular Euler, except the computation of y_{i+1} and x_{i+1}.

$$k_1 = f(x_i, y_i), \qquad x_{i+1} = x_i + h,$$

$$k_2 = f(x_i + h, y_i + k_1 h), \qquad y_{i+1} = y_i + \frac{k_1 + k_2}{2}\, h.$$

Exercise 148 *Consider $\frac{dy}{dx} = x + y$, $y(0) = 1$.*

a) *Use the improved Euler's method (see above) with step sizes $h = 1/4$ and $h = 1/8$ to approximate $y(1)$.*

b) *Use Euler's method with $h = 1/4$ and $h = 1/8$.*

c) *Solve exactly, find the exact value of $y(1)$.*

d) *Compute the errors, and the factors by which the errors changed.*

The simplest method used in practice is the *Runge–Kutta method*. Consider $\frac{dy}{dx} = f(x, y)$, $y(x_0) = y_0$, and a step size h. Everything is the same as in Euler's method, except the computation of y_{i+1} and x_{i+1}.

$$\begin{aligned} k_1 &= f(x_i, y_i), \\ k_2 &= f\big(x_i + h/2, y_i + k_1 (h/2)\big), \qquad & x_{i+1} &= x_i + h, \\ k_3 &= f\big(x_i + h/2, y_i + k_2 (h/2)\big), \qquad & y_{i+1} &= y_i + \frac{k_1 + 2k_2 + 2k_3 + k_4}{6}\, h, \\ k_4 &= f(x_i + h, y_i + k_3 h). \end{aligned}$$

Exercise 149 *Consider $\frac{dy}{dx} = yx^2$, $y(0) = 1$.*

a) *Use Runge–Kutta (see above) with step sizes $h = 1$ and $h = 1/2$ to approximate $y(1)$.*

b) *Use Euler's method with $h = 1$ and $h = 1/2$.*

c) *Solve exactly, find the exact value of $y(1)$, and compare.*

Autonomous equations

We discuss autonomous equations

Definition 7. *An equation of the form*

$$\frac{dx}{dt} = f(x), \tag{10}$$

where the derivative of solutions depends only on x (the dependent variable) is called an autonomous equation*. If we think of t as time, the naming comes from the fact that the equation is independent of time.*

We return to the cooling coffee problem (**Normally a reference to a previous example goes here.**). Newton's law of cooling says

$$\frac{dx}{dt} = k(A - x),$$

where x is the temperature, t is time, k is some positive constant, and A is the ambient temperature. See **Normally a reference to a previous figure goes here.** for an example with $k = 0.3$ and $A = 5$.

Note the solution $x(t) = A$ (in the figure $x = 5$). We call these constant solutions the *equilibrium solutions*. The points on the x-axis where $f(x) = 0$ are called *critical points* of the differential equation (10). The point $x = A$ is a critical point. In fact, each critical point corresponds to an equilibrium solution.

Now, we want to determine what happens for other values of x that are not A. Based on the existence and uniqueness theorem in § for first order differential equations, the fact that $k(A-x)$ and its partial derivative in x, $-k$, are continuous everywhere gives that solution curves can not cross. This means that since we know $x(t) = A$ is a solution, if a solution starts below $x(t) = A$, it must always stay there, and solutions that start above $x(t) = A$ will also stay there. For more information about what the solutions do, we'll need to look back at the equation and some sample solution curves.

Note also, by looking at the graph, that the solution $x = A$ is "stable" in that small perturbations in x do not lead to substantially different solutions as t grows. If we change the initial condition a little bit, then as $t \to \infty$ we get $x(t) \to A$. We call such a critical point *asymptotically stable*. In this simple example, it turns out that all solutions in fact go to A as $t \to \infty$. If there is a critical point where all nearby solutions move away from the critical point, we say it is *unstable*. If some nearby solutions go towards the critical point, and some others move away, then we say it is *semistable*. The final option is that solutions nearby neither move towards nor away from the critical point, and these critical points are called *stable*.

The last of these options may seem strange at first, and that is because stable critical points are not possible for autonomous equations with one unknown function. If a solution does not move towards or away from a critical point, that means it doesn't move anywhere, and so is a critical point on its own. However, when we get to autonomous systems in § **??** and § **??**, we will see that in two dimensions, this is possible (think of a circle that does not spiral into or away from the center point).

Consider now the *logistic equation*

$$\frac{dx}{dt} = kx(M - x),$$

for some positive k and M. This equation is commonly used to model population if we know the limiting population M, that is the maximum sustainable population. The logistic equation leads to less catastrophic predictions on world population than $x' = kx$. In the real world there is no such thing as negative population, but we will still consider negative x for the purposes of the math.

See **Normally a reference to a previous figure goes here.** for an example, $x' = 0.1x(5 - x)$. There are two critical points, $x = 0$ and $x = 5$. The critical point at $x = 5$ is asymptotically stable, while the critical point at $x = 0$ is unstable.

It is not necessary to find the exact solutions to talk about the long term behavior of the solutions. From the slope field above of $x' = 0.1x(5 - x)$, we see that

Learning outcomes: Identify autonomous first order differential equations Find critical points or equilibrium solutions for autonomous equations Sketch a phase line for these equations.

Author(s): Matthew Charnley and Jason Nowell

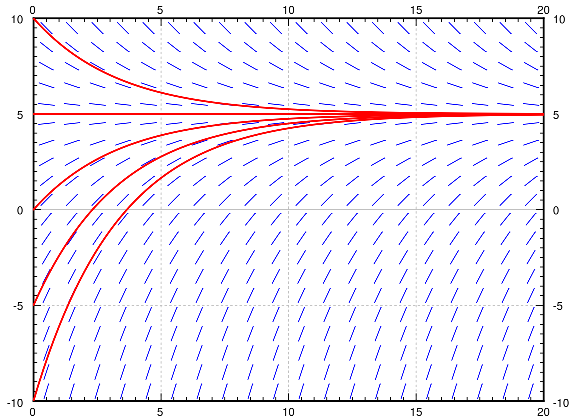


Figure 18: The slope field and some solutions of $x' = 0.3\,(5 - x)$.

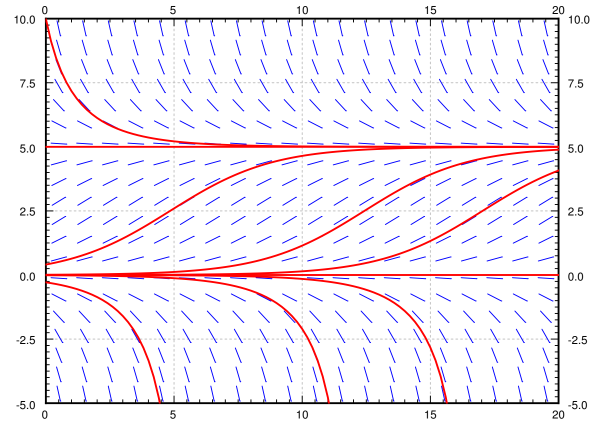


Figure 19: The slope field and some solutions of $x' = 0.1\,x\,(5 - x)$.

$$\lim_{t\to\infty} x(t) = \begin{cases} 5 & \text{if } \; x(0) > 0, \\ 0 & \text{if } \; x(0) = 0, \\ \text{DNE or } -\infty & \text{if } \; x(0) < 0. \end{cases}$$

Here DNE means "does not exist". From just looking at the slope field we cannot quite decide what happens if $x(0) < 0$. It could be that the solution does not exist for t all the way to ∞. Think of the equation $x' = x^2$; we have seen that solutions only exist for some finite period of time. Same can happen here. In our example equation above it turns out that the solution does not exist for all time, but to see that we would have to solve the equation. In any case, the solution does go to $-\infty$, but it may get there rather quickly.

If we are interested only in the long term behavior of the solution, we would be doing unnecessary work if we solved the equation exactly. We could draw the slope field, but it is easier to just look at the *phase diagram* or *phase line*, which is a simple way to visualize the behavior of autonomous equations. The phase line for this equation is visible in **Normally a reference to a previous figure goes here.**. In this case there is one dependent variable x. We draw the x-axis, we mark all the critical points, and then we draw arrows in between. Since x is the dependent variable we draw the axis vertically, as it appears in the slope field diagrams above. If $f(x) > 0$, we draw an up arrow. If $f(x) < 0$, we draw a down arrow. To figure this out, we could just plug in some x between the critical points, $f(x)$ will have the same sign at all x between two critical points as long $f(x)$ is continuous. For example, $f(6) = -0.6 < 0$, so $f(x) < 0$ for $x > 5$, and the arrow above $x = 5$ is a down arrow. Next, $f(1) = 0.4 > 0$, so $f(x) > 0$ whenever $0 < x < 5$, and the arrow points up. Finally, $f(-1) = -0.6 < 0$ so $f(x) < 0$ when $x < 0$, and the arrow points down.

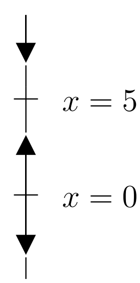


Figure 20: Phase line for the differential equation $x' = 0.1x(5 - x)$.

Armed with the phase diagram, it is easy to sketch the solutions approximately: As time t moves from left to right, the graph of a solution goes up if the arrow is up, and it goes down if the arrow is down.

Exercise 150 *Try sketching a few solutions simply from looking at the phase diagram. Check with the preceding graphs to see if you are getting the types of curves that match the solutions.*

Once we draw the phase diagram, we can use it to classify critical points as asymptotically stable, semistable, or unstable based on whether the "arrows" point into or away from the critical point on each side. Two arrows in means that the critical point is asymptotically stable, two arrows away means unstable, and one in one out means semistable.

Example 21. *Consider the autonomous differential equation*

$$\frac{dx}{dt} = x(x-2)^2(x+3)(x-4) \tag{11}$$

Find all equilibrium solutions for this equation, and determine their stability. Draw a phase line and use this information to sketch some approximate solution curves.

Solution: This equation is already in factored form. This makes it simple to determine the equilibrium solutions as $x = 0$, $x = 2$, $x = -3$ and $x = 4$. In order to determine the stability of each critical point and draw the phase line, we need to plug in values surrounding these points to $f(x) = x(x-2)^2(x+3)(x-4)$. We can see that

$$\begin{aligned} f(-4) &= (-4)(-6)^2(-1)(-8) < 0, \\ f(-1) &= (-1)(-3)^2(2)(-5) > 0, \\ f(1) &= (1)(-1)^2(4)(-3) < 0, \\ f(3) &= (3)(1)^2(6)(-1) < 0, \\ f(5) &= (5)(3)^2(8)(1) > 0. \end{aligned}$$

This lets us draw the phase line and determine the stability of each critical point. Thus, we see that $x = -3$ is an unstable critical point, $x = 0$ is asymptotically stable, $x = 2$ is semistable, and $x = 4$ is unstable. A set of sample solution curves also validates these conclusions.

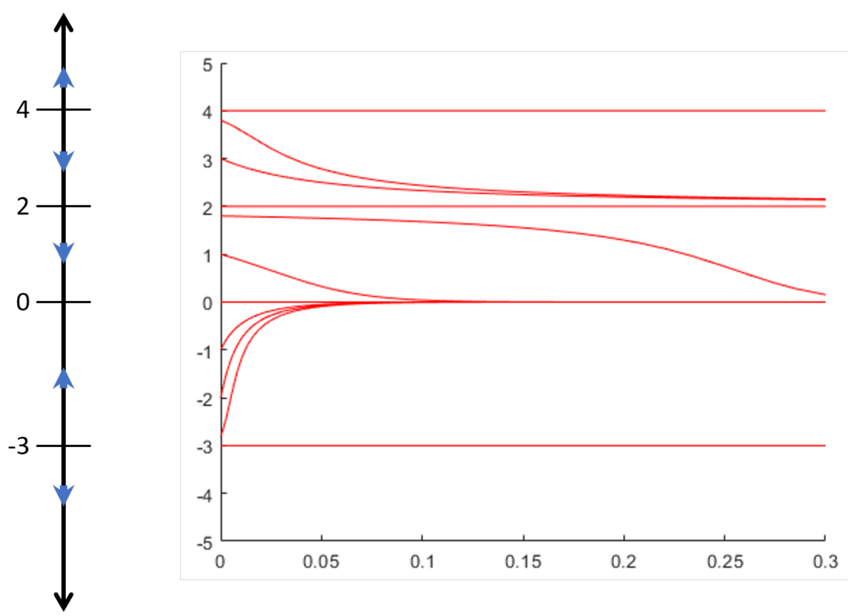


Figure 21: Phase line for the differential equation $\frac{dx}{dt} = x(x-2)^2(x+3)(x-4)$ and a plot of some solutions to this equation.

Concavity of Solutions

We can tell from the phase line for an autonomous equation when the solution will be increasing or decreasing. Is there any more we can learn about the shape of these graphs? There is, and it comes from looking for the concavity, which is determined by the second derivative.

We can compute the second derivative

$$\frac{d^2x}{dt^2} = \frac{d}{dx}\left[\frac{dx}{dt}\right]$$

of our solution by noticing that $\frac{dx}{dt} = f(x)$. This function can be differentiated by the chain rule

$$\frac{d}{dt}f(x) = f'(x)\frac{dx}{dt} = f'(x)f(x).$$

So, the solution is concave up if $f'(x)f(x)$ is positive, and concave down if that is negative. Phrased another way, the solution is concave up if f and f' have the same sign, and it is concave down if f and f' have opposite signs.

Let's see what this looks like in action. Take the logistic equation $x' = 0.1x(5 - x)$, whose solutions are plotted in Figure 19. **Normally a reference to a previous figure goes here.** shows the graph of $f(x)$ as a function of x for this scenario. When do f and f' have the same sign? Well, this happens when f is both positive and increasing, or negative and decreasing. This happens between 0 and the vertex, as well as for $x > 5$. The vertex here is at $x = 2.5$, and so we conclude that the solution should be concave up when x is on the intervals $(0, 2.5)$ and $(5, \infty)$, and be concave down otherwise. Looking back at Figure 19, this is exactly what we observe. All of the solutions between 0 and 5 seem to "flip over" to be concave down when x crosses 2.5.

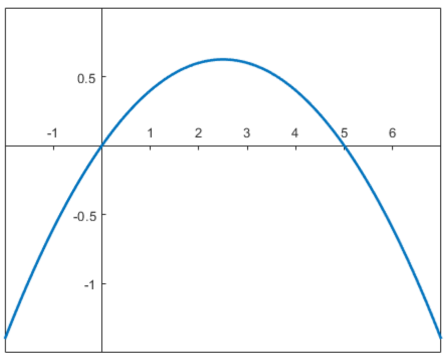


Figure 22: Plot of x vs. $f(x)$ for the differential equation $\frac{dx}{dt} = 0.1x(5 - x)$.

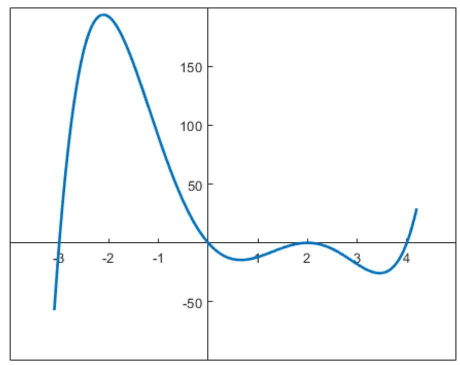


Figure 23: Plot of x vs. $f(x)$ for the differential equation $\frac{dx}{dt} = x(x-2)^2(x+3)(x-4)$.

The same can be seen for solutions to (11), even though we can't compute the extreme values explicitly. **Normally a reference to a previous figure goes here.** shows the graph of $f(x)$ vs. x for this situation. Between each pair of equilibrium solutions there is a critical point of f (in the Calculus 1 sense) where the derivative is zero, and at this point, the derivative changes sign, and since the function value does not change sign, the concavity of the solution to the differential equation flips at this point. Comparing this graph and these points where concavity shifts with the solutions drawn in Figure 21 again validates these results.

firstOrder/practice/autonomous-eqn-practice1.tex

Practice for Autonomous Equations

Why?

Exercise 151 *Consider* $x' = x^2$.

a) *Draw the phase diagram, find the critical points, and mark them asymptotically stable, semistable, or unstable.*

b) *Sketch typical solutions of the equation.*

c) *Find* $\lim_{t\to\infty} x(t)$ *for the solution with the initial condition* $x(0) = -1$.

Exercise 152 *Consider* $x' = \sin x$.

a) *Draw the phase diagram for* $-4\pi \leq x \leq 4\pi$. *On this interval mark the critical points asymptotically stable, semistable, or unstable.*

b) *Sketch typical solutions of the equation.*

c) *Find* $\lim_{t\to\infty} x(t)$ *for the solution with the initial condition* $x(0) = 1$.

Exercise 153 *Let* $x' = (x-1)(x-2)x^2$.

a) *Sketch the phase diagram and find critical points.*

b) *Classify the critical points.*

c) *If* $x(0) = 0.5$, *then find* $\lim_{t\to\infty} x(t)$.

Exercise 154 *Let* $y' = (y-2)(y^2+1)(y+3)$. *Sketch a phase diagram for this differential equation. Find and classify all critical points. If* $y(0) = 0$, *what will happen to the solution as* $t \to \infty$?

Exercise 155 *Find and classify all equilibrium solutions for the differential equation* $x' = (x-2)^2(x+1)(x+3)^3(x+2)$.

Exercise 156 *Let* $y' = (y-3)(y+2)^2 e^y$. *Sketch a phase diagram for this differential equation. Find and classify all critical points. If* $y(0) = 0$, *what will happen to the solution as* $t \to \infty$?

Exercise 157 *Consider the DE* $\dfrac{dy}{dt} = y^5 - 3y^4 + 3y^3 - y^2$. *Find and classify all equilibrium solutions of this DE. Then sketch a representative selection of solution curves.*

Author(s): Matthew Charnley and Jason Nowell

Exercise 158 *Let* $x' = e^{-x}$.

a) *Find and classify all critical points.*

b) *Find* $\lim_{t\to\infty} x(t)$ *given any initial condition.*

Exercise 159 *Suppose* $f(x)$ *is positive for* $0 < x < 1$, *it is zero when* $x = 0$ *and* $x = 1$, *and it is negative for all other* x.

a) *Draw the phase diagram for* $x' = f(x)$, *find the critical points, and mark them asymptotically stable, semistable, or unstable.*

b) *Sketch typical solutions of the equation.*

c) *Find* $\lim_{t\to\infty} x(t)$ *for the solution with the initial condition* $x(0) = 0.5$.

Exercise 160 *Suppose* $\frac{dx}{dt} = (x - \alpha)(x - \beta)$ *for two numbers* $\alpha < \beta$.

a) *Find the critical points, and classify them.*

For b), c), d), find $\lim_{t\to\infty} x(t)$ *based on the phase diagram.*

b) $x(0) < \alpha$,

c) $\alpha < x(0) < \beta$,

d) $\beta < x(0)$.

Exercise 161 *A disease is spreading through the country. Let* x *be the number of people infected. Let the constant* S *be the number of people susceptible to infection. The infection rate* $\frac{dx}{dt}$ *is proportional to the product of already infected people,* x, *and the number of susceptible but uninfected people,* $S - x$.

a) *Write down the differential equation.*

b) *Supposing* $x(0) > 0$, *that is, some people are infected at time* $t = 0$, *what is* $\lim_{t\to\infty} x(t)$.

c) *Does the solution to part b) agree with your intuition? Why or why not?*

Bifurcation diagrams

We discuss bifurcation diagrams for autonomous equations with parameters.

An extension of the topic of autonomous equation is *autonomous equations with parameter.* The idea is that we have a differential equation that has no explicit dependence on time, but does have a dependence on an outside parameter, which is a constant set by the physical situation. In terms of physical problems, this parameter will tend to be something that we can adjust to change how the differential equation behaves. For example, in a logistic differential equation

$$\frac{dx}{dt} = ax(K - x)$$

either the a or the K (or both) could be adjustable parameters. For a given value of the parameter, the differential equation behaves like a standard autonomous differential equation, but we can get different properties of this equation for different values of the parameter.

Definition 8. *m An* autonomous equation with parameter α *is a differential equation of the form*

$$\frac{dx}{dt} = f_\alpha(x)$$

where, for every value of α, $f_\alpha(x)$ is a function of the single variable x.

Later, we will want to view $f_\alpha(x)$ as a two-variable function of x and α, but for now, we want to think about it as a function of just x for a fixed value of α. We want to be able to analyze what happens to this equation for different values of α. Since it is an autonomous equation, we can do this using phase lines. This will be easiest to see through an example.

Example 22. *Consider the differential equation*

$$\frac{dx}{dt} = x(x^2 - \alpha),$$

which fits the description of an autonomous equation with parameter α. Describe what happens in this differential equation for $\alpha = -4$, $\alpha = 0$, and $\alpha = 1$.

Solution: We can draw a phase line for $\alpha = -4$, $\alpha = 0$ and $\alpha = 1$. It is clear that something happens with this equation between $\alpha = -4$ and $\alpha = 1$. We go from having only one equilibrium solution at $\alpha = -4$ to having three equilibrium solutions at $\alpha = 1$. In addition, the solution at $y = 0$ is unstable for $\alpha = -4$, while it is asymptotically stable for $\alpha = 1$. If we want to figure out when this change happens, we'll need a better way to analyze this problem.

How can we better approach this problem? The idea is to think about when the solution to the differential equation will be increasing or decreasing as a function of the two variables α and x. Based on the structure of the differential equation, the solution will be increasing when the function $f_\alpha(x)$ is positive and will be decreasing when $f_\alpha(x)$ is negative. Since a phase line is a plot of this information for a given value of α, we essentially want to plot all of these phase lines on a two-dimensional graph. This graph is called a *bifurcation diagram.* **Normally a reference to a previous figure goes here.** shows a bifurcation diagram for the example $\frac{dx}{dt} = x(x^2 - \alpha)$.

Within this picture, we can see all of our phase lines from before, because at any value of α, taking the vertical slice of this graph at that value, we get the phase line. If we want to consider $\alpha = -4$, then we can look above the horizontal coordinate -4, and that will give us the phase line for $\alpha = -4$. The same goes for any other value of α we want to consider. For instance, we can also see that for any $\alpha \leq 0$, there will be one equilibrium solution, and for $\alpha > 0$ there are three equilibrium solutions, indicated by the three black curves above each of those α values.

From this, we can see that the point at which the behavior changes is $\alpha = 0$. Thus, for this problem $\alpha = 0$ is called the *bifurcation point.* This is defined to be the value of the parameter for which the overall behavior of the equation changes. This can be a change in the number of equilibrium solutions, the stability of these equilibrium solutions, or both. For this particular example, we have both of these. We go from 1 equilibrium solution to 3, and the solution at $y = 0$ changes in

Learning outcomes: Draw and analyze bifurcation diagrams for autonomous equations with parameter.

Author(s): Matthew Charnley and Jason Nowell

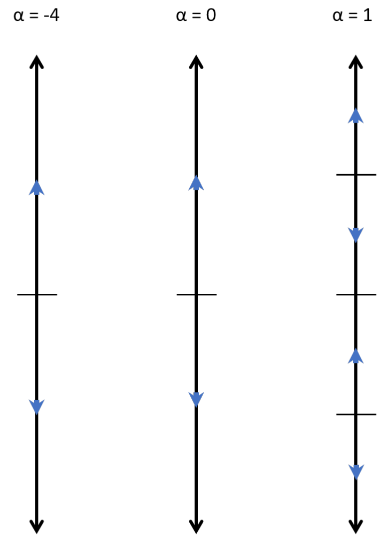


Figure 24: Phase lines for the differential equation $\frac{dx}{dt} = x(x^2 - \alpha)$ for $\alpha = -4,\ 0,\ 1$.

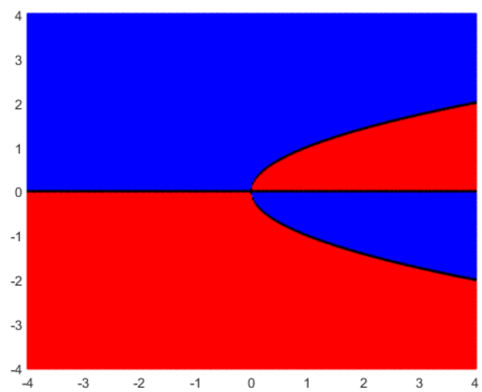


Figure 25: Bifurcation Diagram for the differential equation $\frac{dx}{dt} = x(x^2 - \alpha)$. In this figure, a blue region means the solution will be increasing and red indicates decreasing.

stability. This type of bifurcation is called a "pitchfork bifurcation" based on the shape of the equilibrium solutions near the bifurcation point.

Another example of a bifurcation of a different form can be seen in the example of the logistic equation with harvesting. Suppose an alien race really likes to eat humans. They keep a planet with humans on it and harvest the humans at a rate of h million humans per year. Suppose x is the number of humans in millions on the planet and t is time in years. Let M be the limiting population when no harvesting is done. The number $k > 0$ is a constant depending on how fast humans multiply. Our equation becomes

$$\frac{dx}{dt} = kx(M - x) - h.$$

In this setup, M and k are fixed values, and the parameter that is being adjusted for this equation is h. We expand the right-hand side and set it to zero.

$$kx(M - x) - h = -kx^2 + kMx - h = 0.$$

Solving for the critical points using the quadratic formula, let us call them A and B, we get

$$A = \frac{kM + \sqrt{(kM)^2 - 4hk}}{2k}, \qquad B = \frac{kM - \sqrt{(kM)^2 - 4hk}}{2k}.$$

Exercise 162 *Sketch a phase diagram for different possibilities. Note that these possibilities are $A > B$, or $A = B$, or A and B both complex (i.e. no real solutions). Hint: Fix some simple k and M and then vary h.*

Example 23. *For example, let $M = 8$ and $k = 0.1$. What happens for different values of h in this situation?*

Solution: When $h = 1$, then A and B are distinct and positive. The slope field we get is in **Normally a reference to a previous figure goes here.**. As long as the population starts above B, which is approximately 1.55 million, then the population will not die out. It will in fact tend towards $A \approx 6.45$ million. If ever some catastrophe happens and the population drops below B, humans will die out, and the fast food restaurant serving them will go out of business.

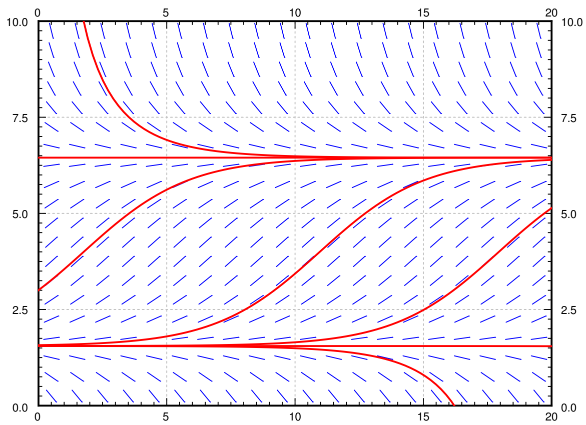


Figure 26: The slope field and some solutions of $x' = 0.1\,x\,(8 - x) - 1$.

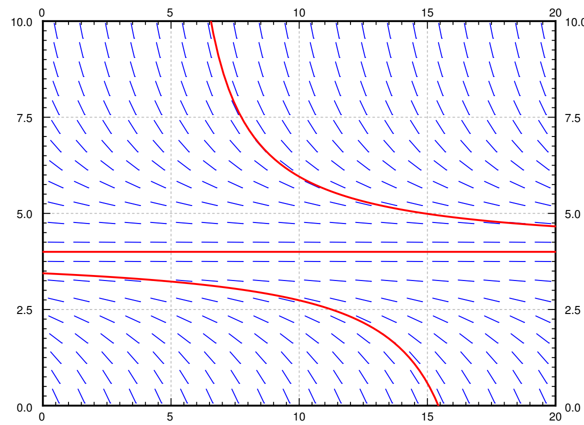


Figure 27: The slope field and some solutions of $x' = 0.1\,x\,(8 - x) - 1.6$.

When $h = 1.6$, then $A = B = 4$. There is only one critical point and it is semistable. When the population starts above 4 million it will tend towards 4 million. If it ever drops below 4 million, humans will die out on the planet. This scenario is not one that we (as the human fast food proprietor) want to be in. A small perturbation of the equilibrium state and we are out of business. There is no room for error. See **Normally a reference to a previous figure goes here.**.

Finally if we are harvesting at 2 million humans per year, there are no critical points. The population will always plummet towards zero, no matter how well stocked the planet starts. See **Normally a reference to a previous figure goes here.**.

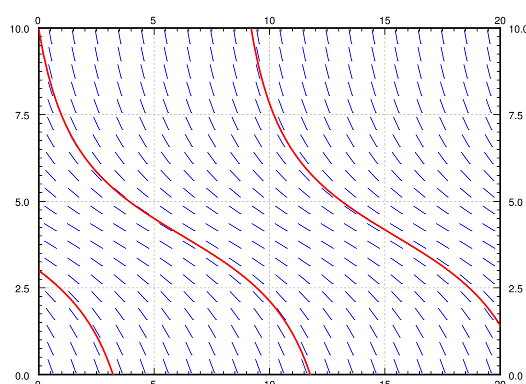


Figure 28: The slope field and some solutions of $x' = 0.1\,x\,(8 - x) - 2$.

All of these can also be seen from the bifurcation diagram, which is drawn in **Normally a reference to a previous figure goes here.**. The values A and B discussed above represent the upper and lower branches of the parabola in the figure. For any $h > 1.6$, there are no equilibrium solutions and the phase line is entirely decreasing, meaning the solution will converge to zero no matter what. For $h < 1.6$, there are two equilibrium solutions, with the top one asymptotically stable and the bottom one unstable. At $h = 1.6$ is where the bifurcation point occurs for this example. This is an example of a "saddle-node" bifurcation, as the two equilibrium solutions collide with each other at the bifurcation point and disappear.

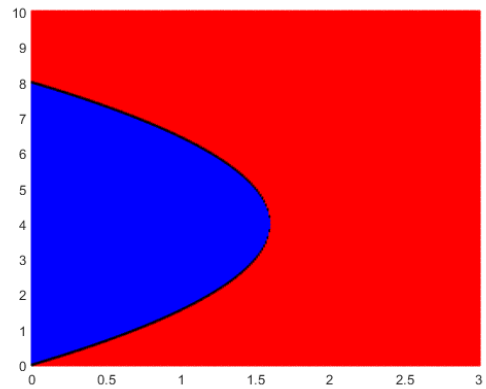


Figure 29: Bifurcation diagram for the differential equation $x' = 0.1\,x\,(8-x) - h$.

Another way to visualize this situation is by plotting the function $f_\alpha(x)$ for the different values of α. The places where this function is zero give the equilibrium solutions, and we can determine *bifurcation values* by looking for where the zeros of this function change behavior. For this particular example, the graphs of $f_\alpha(x)$ are drawn in **Normally a reference to a previous figure goes here.**.

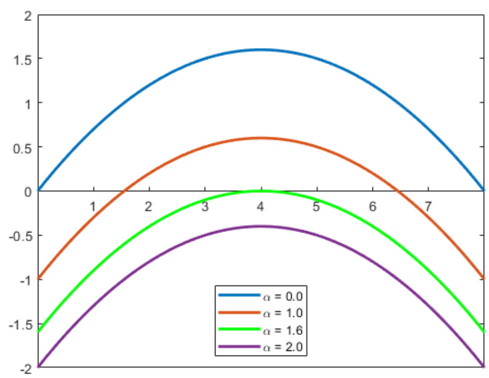


Figure 30: Graph of $f_\alpha(x) = 0.1x(8-x) - \alpha$ for $\alpha = 0$, 1.0, 1.6, 2.0.

The values of α we are looking for are those where the number and types of zeros change for the function $f_\alpha(x)$. In this figure, we see that for $\alpha < 1.6$, the parabola crosses the x axis twice, resulting in two zeros and two equilibrium solutions. For $\alpha = 1.6$, there is one (double) root, and for $\alpha > 1.6$, there are no equilibrium solutions, and the function $f_\alpha(x)$ is always negative. Since the number of roots/zeros changes at $\alpha = 1.6$, that means that 1.6 is the bifurcation point for this equation. We can also see this from the equation, since the equilibrium solutions are determined by the values of x where

$$0.1x(8-x) - \alpha = 0 \qquad \text{or} \qquad -0.1x^2 + 0.8x - \alpha = 0$$

which can be found by the quadratic formula

$$x = \frac{0.8 \pm \sqrt{0.64 - 4(0.1)(\alpha)}}{0.2}.$$

Roots to this equation do not exist (because they are complex) if $0.64 - 0.4\alpha < 0$, or $\alpha > 1.6$.

firstOrder/practice/bifurcation-diag-practice1.tex

Practice for bifurcation diagrams

Why?

Exercise 163 *Start with the logistic equation* $\frac{dx}{dt} = kx(M - x)$*. Suppose we modify our harvesting. That is we will only harvest an amount proportional to current population. In other words, we harvest* hx *per unit of time for some* $h > 0$ *(Similar to earlier example with* h *replaced with* hx*).*

a) *Construct the differential equation.*

b) *Show that if* $kM > h$*, then the equation is still logistic.*

c) *What happens when* $kM < h$*?*

Exercise 164 *Assume that a population of fish in a lake satisfies* $\frac{dx}{dt} = kx(M-x)$*. Now suppose that fish are continually added at* A *fish per unit of time.*

a) *Find the differential equation for* x*.*

b) *What is the new limiting population?*

Exercise 165 *Consider the differential equation with parameter* α *given by* $y' = y(y - \alpha + 1)$*.*

a) *Sketch a phase diagram for this differential equation with* $\alpha = -3$*,* $\alpha = 1$*, and* $\alpha = 3$*.*

b) *Draw a bifurcation diagram for this differential equation with parameter.*

c) *What is the bifurcation point for this equation? What changes when* α *passes over the bifurcation point?*

Exercise 166 *Consider the differential equation with parameter* α *given by* $y' = y^2(y^2 - \alpha)$*.*

a) *Sketch a phase diagram for this differential equation with* $\alpha = -3$*,* $\alpha = 0$*, and* $\alpha = 3$*.*

b) *Draw a bifurcation diagram for this differential equation with parameter.*

c) *What is the bifurcation point for this equation? What changes when* α *passes over the bifurcation point?*

Exercise 167 *Consider the differential equation with parameter* α *given by* $y' = y(\alpha - y)$*.*

a) *Sketch a phase diagram for this differential equation with* $\alpha = -3$*,* $\alpha = 0$*, and* $\alpha = 3$*.*

b) *Draw a bifurcation diagram for this differential equation with parameter.*

c) *What is the bifurcation point for this equation? What changes when* α *passes over the bifurcation point?*

Author(s): Matthew Charnley and Jason Nowell

s

Exact equations

We introduce how to classify various properties of differential equations.

Another type of equation that comes up quite often in physics and engineering is an *exact equation.* Suppose $F(x,y)$ is a function of two variables, which we call the *potential function.* The naming should suggest potential energy, or electric potential. Exact equations and potential functions appear when there is a conservation law at play, such as conservation of energy. Let us make up a simple example. Let

$$F(x,y) = x^2 + y^2.$$

We are interested in the lines of constant energy, that is lines where the energy is conserved; we want curves where $F(x,y) = C$, for some constant C, since F represents the energy of the system. In our example, the curves $x^2+y^2 = C$ are circles. See **Normally a reference to a previous figure goes here.**.

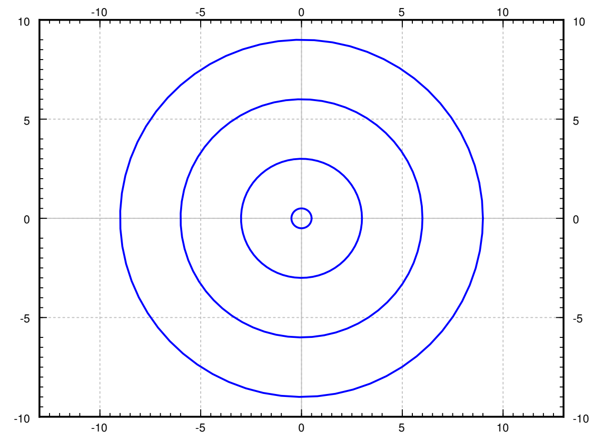


Figure 31: Solutions to $F(x,y) = x^2 + y^2 = C$ for various C.

We take the *total derivative* of F:

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy.$$

For convenience, we will make use of the notation of $F_x = \frac{\partial F}{\partial x}$ and $F_y = \frac{\partial F}{\partial y}$. In our example,

$$dF = 2x\,dx + 2y\,dy.$$

We apply the total derivative to $F(x,y) = C$, to find the differential equation $dF = 0$. The differential equation we obtain in such a way has the form

$$M\,dx + N\,dy = 0, \qquad \text{or} \qquad M + N\,\frac{dy}{dx} = 0.$$

Definition 9. *An equation of the form*

$$M(x,y) + N(x,y)\,\frac{dy}{dx} = 0$$

is called exact *if it was obtained as $dF = 0$ for some potential function F.*

In our simple example, we obtain the equation

$$2x\,dx + 2y\,dy = 0, \qquad \text{or} \qquad 2x + 2y\,\frac{dy}{dx} = 0.$$

Since we obtained this equation by differentiating $x^2 + y^2 = C$, the equation is exact. We often wish to solve for y in terms of x. In our example,

$$y = \pm\sqrt{C^2 - x^2}.$$

An interpretation of the setup is that at each point in the plane $\vec{v} = (M,N)$ is a vector, that is, a direction and a magnitude. As M and N are functions of (x,y), we have a *vector field.* The particular vector field $\vec{v}$ that comes from an exact equation is a so-called *conservative vector field*, that is, a vector field that comes with a potential function $F(x,y)$, such that

$$\vec{v} = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}\right).$$

Learning outcomes: Determine if a first order differential equation is exact Find the general solution to an exact equation Solve initial value problems for exact equations Use integrating factors to make some non-exact equations exact in order to solve them.

Author(s): Matthew Charnley and Jason Nowell

This is something that you may have seen in your Calculus 3 course, and if so, the process for solving exact equations is basically identical to the process of finding a potential function for a conservative vector field. The physical interpretation of conservative vector fields is as follows. Let γ be a path in the plane starting at (x_1, y_1) and ending at (x_2, y_2). If we think of $\vec{v}$ as force, then the work required to move along γ is

$$\int_\gamma \vec{v}(\vec{r}) \cdot d\vec{r} = \int_\gamma M\,dx + N\,dy = F(x_2, y_2) - F(x_1, y_1).$$

That is, the work done only depends on endpoints, that is where we start and where we end. For example, suppose F is gravitational potential. The derivative of F given by $\vec{v}$ is the gravitational force. What we are saying is that the work required to move a heavy box from the ground floor to the roof only depends on the change in potential energy. That is, the work done is the same no matter what path we took; if we took the stairs or the elevator. Although if we took the elevator, the elevator is doing the work for us. The curves $F(x, y) = C$ are those where no work need be done, such as the heavy box sliding along without accelerating or breaking on a perfectly flat roof, on a cart with incredibly well oiled wheels. Effectively, an exact equation is a conservative vector field, and the implicit solution of this equation is the potential function.

Solving exact equations

Now you, the reader, should ask: Where did we solve a differential equation? Well, in applications we generally know M and N, but we do not know F. That is, we may have just started with $2x + 2y\frac{dy}{dx} = 0$, or perhaps even

$$x + y\frac{dy}{dx} = 0.$$

It is up to us to find some potential F that works. Many different F will work; adding a constant to F does not change the equation. Once we have a potential function F, the equation $F\big(x, y(x)\big) = C$ gives an implicit solution of the ODE.

Example 24. *Let us find the general solution to $2x + 2y\frac{dy}{dx} = 0$. Forget we knew what F was.*

Solution: If we know that this is an exact equation, we start looking for a potential function F. We have $M = 2x$ and $N = 2y$. If F exists, it must be such that $F_x(x, y) = 2x$. Integrate in the x variable to find

$$F(x, y) = x^2 + A(y), \tag{12}$$

for some function $A(y)$. The function A is the "constant of integration", though it is only constant as far as x is concerned, and may still depend on y. Now differentiate (12) in y and set it equal to N, which is what F_y is supposed to be:

$$2y = F_y(x, y) = A'(y).$$

Integrating, we find $A(y) = y^2$. We could add a constant of integration if we wanted to, but there is no need. We found $F(x, y) = x^2 + y^2$. Next for a constant C, we solve

$$F\big(x, y(x)\big) = C.$$

for y in terms of x. In this case, we obtain $y = \pm\sqrt{C^2 - x^2}$ as we did before.

Exercise 168 *Why did we not need to add a constant of integration when integrating $A'(y) = 2y$? Add a constant of integration, say 3, and see what F you get. What is the difference from what we got above, and why does it not matter?*

In the previous example, you may have also noticed that the equation $2x + 2y\frac{dy}{dx} = 0$ is separable, and we could have solved it via that method as well. This is not a coincidence, as every separable equation is exact (see Exercise for the details) but there are many exact equations that are not separable, which we will see throughout the examples here.

The procedure, once we know that the equation is exact, is:

(a) Integrate $F_x = M$ in x resulting in $F(x, y) = \text{something} + A(y)$.

(b) Differentiate this F in y, and set that equal to N, so that we may find $A(y)$ by integration.

The procedure can also be done by first integrating in y and then differentiating in x. Pretty easy huh? Let's try this again.

Example 25. *Consider now* $2x + y + xy\frac{dy}{dx} = 0$.

Solution: OK, so $M = 2x + y$ and $N = xy$. We try to proceed as before. Suppose F exists. Then $F_x(x,y) = 2x + y$. We integrate:

$$F(x,y) = x^2 + xy + A(y)$$

for some function $A(y)$. Differentiate in y and set equal to N:

$$N = xy = F_y(x,y) = x + A'(y).$$

But there is no way to satisfy this requirement! The function xy cannot be written as x plus a function of y. The equation is not exact; no potential function F exists.

Is there an easier way to check for the existence of F, other than failing in trying to find it? Turns out there is. Suppose $M = F_x$ and $N = F_y$. Then as long as the second derivatives are continuous,

$$\frac{\partial M}{\partial y} = \frac{\partial^2 F}{\partial y \partial x} = \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial N}{\partial x}.$$

Let us state it as a theorem. Usually this is called the Poincaré Lemma[1].

Theorem 3 (thm:Poincare)**.** *Poincaré If M and N are continuously differentiable functions of (x,y), and $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$, then near any point there is a function $F(x,y)$ such that $M = \frac{\partial F}{\partial x}$ and $N = \frac{\partial F}{\partial y}$.*

The theorem doesn't give us a global F defined everywhere. In general, we can only find the potential locally, near some initial point. By this time, we have come to expect this from differential equations.

Let us return to the example above where $M = 2x + y$ and $N = xy$. Notice $M_y = 1$ and $N_x = y$, which are clearly not equal. The equation is not exact.

Example 26. *Solve*

$$\frac{dy}{dx} = \frac{-2x - y}{x - 1}, \qquad y(0) = 1.$$

Solution: We write the equation as

$$(2x + y) + (x - 1)\frac{dy}{dx} = 0,$$

so $M = 2x + y$ and $N = x - 1$. Then

$$M_y = 1 = N_x.$$

The equation is exact. Integrating M in x, we find

$$F(x,y) = x^2 + xy + A(y).$$

Differentiating in y and setting to N, we find

$$x - 1 = x + A'(y).$$

So $A'(y) = -1$, and $A(y) = -y$ will work. Take $F(x,y) = x^2 + xy - y$. We wish to solve $x^2 + xy - y = C$. First let us find C. As $y(0) = 1$ then $F(0,1) = C$. Therefore $0^2 + 0 \times 1 - 1 = C$, so $C = -1$. Now we solve $x^2 + xy - y = -1$ for y to get

$$y = \frac{-x^2 - 1}{x - 1}.$$

[1] Named for the French polymath Jules Henri Poincaré (1854–1912).

Example 27. *Solve*

$$-\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy = 0, \qquad y(1) = 2.$$

Solution: We leave to the reader to check that $M_y = N_x$.

This vector field (M, N) is not conservative if considered as a vector field of the entire plane minus the origin. The problem is that if the curve γ is a circle around the origin, say starting at $(1, 0)$ and ending at $(1, 0)$ going counterclockwise, then if F existed we would expect

$$0 = F(1,0) - F(1,0) = \int_\gamma F_x\,dx + F_y\,dy = \int_\gamma \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy = 2\pi.$$

That is nonsense! We leave the computation of the path integral to the interested reader, or you can consult your multivariable calculus textbook. So there is no potential function F defined everywhere outside the origin $(0, 0)$.

If we think back to the theorem, it does not guarantee such a function anyway. It only guarantees a potential function locally, that is only in some region near the initial point. As $y(1) = 2$ we start at the point $(1, 2)$. Considering $x > 0$ and integrating M in x or N in y, we find

$$F(x, y) = \arctan\left({}^{y}\!/_{x}\right).$$

The implicit solution is $\arctan\left({}^{y}\!/_{x}\right) = C$. Solving, $y = \tan(C)x$. That is, the solution is a straight line. Solving $y(1) = 2$ gives us that $\tan(C) = 2$, and so $y = 2x$ is the desired solution. See **Normally a reference to a previous figure goes here.**, and note that the solution only exists for $x > 0$.

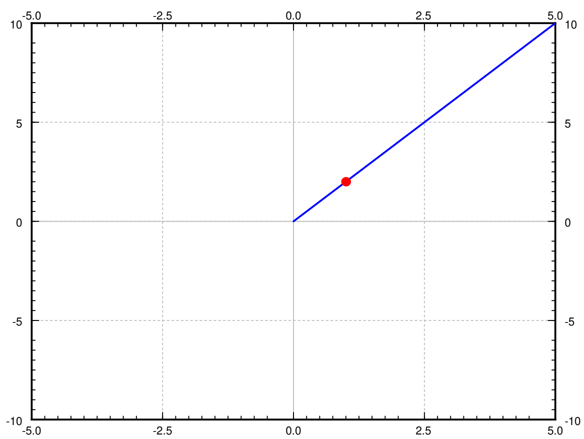


Figure 32: Solution to $-\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy = 0$, $y(1) = 2$, with initial point marked.

Example 28. *Solve*

$$x^2 + y^2 + 2y(x+1)\frac{dy}{dx} = 0.$$

Solution: The reader should check that this equation is exact. Let $M = x^2 + y^2$ and $N = 2y(x + 1)$. We follow the procedure for exact equations

$$F(x, y) = \frac{1}{3}x^3 + xy^2 + A(y),$$

and

$$2y(x + 1) = 2xy + A'(y).$$

Therefore $A'(y) = 2y$ or $A(y) = y^2$ and $F(x, y) = \frac{1}{3}x^3 + xy^2 + y^2$. We try to solve $F(x, y) = C$. We easily solve for y^2 and then just take the square root:

$$y^2 = \frac{C - ({}^{1}\!/_{3})x^3}{x+1}, \qquad \text{so} \qquad y = \pm\sqrt{\frac{C - ({}^{1}\!/_{3})x^3}{x+1}}.$$

When $x = -1$, the term in front of $\frac{dy}{dx}$ vanishes. You can also see that our solution is not valid in that case. However, one could in that case try to solve for x in terms of y starting from the implicit solution $\frac{1}{3}x^3 + xy^2 + y^2 = C$. The solution is somewhat messy and we leave it as implicit.

Integrating factors

Sometimes an equation $M\,dx + N\,dy = 0$ is not exact, but it can be made exact by multiplying with a function $u(x, y)$. That is, perhaps for some nonzero function $u(x, y)$,

$$u(x, y)M(x, y)\,dx + u(x, y)N(x, y)\,dy = 0$$

is exact. Any solution to this new equation is also a solution to $M\,dx + N\,dy = 0$.

In fact, a linear equation

$$\frac{dy}{dx} + p(x)y = f(x), \qquad \text{or} \qquad \bigl(p(x)y - f(x)\bigr)\,dx + dy = 0$$

is always such an equation. Let $r(x) = e^{\int p(x)\,dx}$ be the integrating factor for a linear equation. Multiply the equation by $r(x)$ and write it in the form of $M + N\frac{dy}{dx} = 0$.

$$r(x)p(x)y - r(x)f(x) + r(x)\frac{dy}{dx} = 0.$$

Then $M = r(x)p(x)y - r(x)f(x)$, so $M_y = r(x)p(x)$, while $N = r(x)$, so $N_x = r'(x) = r(x)p(x)$. In other words, we have an exact equation. Integrating factors for linear functions are just a special case of integrating factors for exact equations.

But how do we find the integrating factor u? Well, given an equation

$$M\,dx + N\,dy = 0,$$

u should be a function such that

$$\frac{\partial}{\partial y}\bigl[uM\bigr] = u_y M + u M_y = \frac{\partial}{\partial x}\bigl[uN\bigr] = u_x N + u N_x.$$

Therefore,

$$(M_y - N_x)u = u_x N - u_y M.$$

At first it may seem we replaced one differential equation by another. True, but all hope is not lost.

A strategy that often works is to look for a u that is a function of x alone, or a function of y alone. If u is a function of x alone, that is $u(x)$, then we write $u'(x)$ instead of u_x, and u_y is just zero. Then

$$\frac{M_y - N_x}{N}u = u'.$$

In particular, $\frac{M_y - N_x}{N}$ ought to be a function of x alone (not depend on y). If so, then we have a linear equation

$$u' - \frac{M_y - N_x}{N}u = 0.$$

Letting $p(x) = \frac{M_y - N_x}{N}$, we solve using the standard integrating factor method, to find $u(x) = Ce^{\int p(x)\,dx}$. The constant in the solution is not relevant, we need any nonzero solution, so we take $C = 1$. Then $u(x) = e^{\int p(x)\,dx}$ is the integrating factor.

Similarly we could try a function of the form $u(y)$. Then

$$\frac{M_y - N_x}{M}u = -u'.$$

In particular, $\frac{M_y - N_x}{M}$ ought to be a function of y alone. If so, then we have a linear equation

$$u' + \frac{M_y - N_x}{M}u = 0.$$

Letting $q(y) = \frac{M_y - N_x}{M}$, we find $u(y) = Ce^{-\int q(y)\,dy}$. We take $C = 1$. So $u(y) = e^{-\int q(y)\,dy}$ is the integrating factor.

Example 29. *Solve*

$$\frac{x^2+y^2}{x+1} + 2y\frac{dy}{dx} = 0.$$

Solution: Let $M = \dfrac{x^2+y^2}{x+1}$ and $N = 2y$. Compute

$$M_y - N_x = \frac{2y}{x+1} - 0 = \frac{2y}{x+1}.$$

As this is not zero, the equation is not exact. We notice

$$P(x) = \frac{M_y - N_x}{N} = \frac{2y}{x+1}\frac{1}{2y} = \frac{1}{x+1}$$

is a function of x alone. We compute the integrating factor

$$e^{\int P(x)\,dx} = e^{\ln|x+1|} = |x+1|.$$

Assuming that we want to look at $x > -1$, we multiply our given equation by $(x+1)$ to obtain

$$x^2 + y^2 + 2y(x+1)\frac{dy}{dx} = 0,$$

which is an exact equation that we solved in **Normally a reference to a previous example goes here.**. The solution was

$$y = \pm\sqrt{\frac{C - (1/3)x^3}{x+1}}.$$

If, instead, we had wanted a solution with $x < -1$, we would have needed to multiply by $-(x+1)$, which would have given a very similar result.

Example 30. *Solve*

$$y^2 + (xy+1)\frac{dy}{dx} = 0.$$

Solution: First compute

$$M_y - N_x = 2y - y = y.$$

As this is not zero, the equation is not exact. We observe

$$Q(y) = \frac{M_y - N_x}{M} = \frac{y}{y^2} = \frac{1}{y}$$

is a function of y alone. We compute the integrating factor

$$e^{-\int Q(y)\,dy} = e^{-\ln y} = \frac{1}{y}.$$

Therefore we look at the exact equation

$$y + \frac{xy+1}{y}\frac{dy}{dx} = 0.$$

The reader should double check that this equation is exact. We follow the procedure for exact equations

$$F(x,y) = xy + A(y),$$

and

$$\frac{xy+1}{y} = x + \frac{1}{y} = x + A'(y). \tag{13}$$

Consequently $A'(y) = \dfrac{1}{y}$ or $A(y) = \ln y$. Thus $F(x,y) = xy + \ln y$. It is not possible to solve $F(x,y) = C$ for y in terms of elementary functions, so let us be content with the implicit solution:

$$xy + \ln y = C.$$

We are looking for the general solution and we divided by y above. We should check what happens when $y = 0$, as the equation itself makes perfect sense in that case. We plug in $y = 0$ to find the equation is satisfied. So $y(x) = 0$ is also a solution.

`firstOrder/practice/exactFunc-practice1.tex`

Practice for Solving Exact ODEs

Why?

Exercise 169 *Solve the following exact equations, implicit general solutions will suffice:*

a) $(2xy + x^2)\, dx + (x^2 + y^2 + 1)\, dy = 0$

b) $x^5 + y^5 \dfrac{dy}{dx} = 0$

c) $e^x + y^3 + 3xy^2 \dfrac{dy}{dx} = 0$

d) $(x + y)\cos(x) + \sin(x) + \sin(x)y' = 0$

Exercise 170 *Solve the following exact equations, implicit general solutions will suffice:*

a) $\cos(x) + ye^{xy} + xe^{xy}y' = 0$

b) $(2x + y)\, dx + (x - 4y)\, dy = 0$

c) $e^x + e^y \dfrac{dy}{dx} = 0$

d) $(3x^2 + 3y)\, dx + (3y^2 + 3x)\, dy = 0$

Exercise 171 *Solve the differential equation* $(2ye^{2xy} - 2x) + (2xe^{2xy} + \cos(y))y' = 0$

Exercise 172 *Solve the differential equation* $(-y\sin(xy) - 2xe^{x^2}) + (-x\sin(xy) + 1)y' = 0$

Exercise 173 *Solve the differential equation* $(2x + 3y\sin(xy)) + (3x\sin(xy) - e^y)y' = 0$ *with* $y(2) = 0$.

Exercise 174 *Solve the differential equation* $x + yy' = 0$ *with* $y(0) = 8$. *Write this as an explicit function and determine the interval of* x *values where the solution is valid.*

Exercise 175 *Solve the differential equation* $2x - 2 + (8y + 16)y' = 0$ *with* $y(2) = 0$. *Write this as an explicit function and determine the interval of* x *values where the solution is valid.*

Exercise 176 *Find the integrating factor for the following equations making them into exact equations. You can either use the formulas in this section or guess what the integrating factor should be.*

a) $e^{xy}\, dx + \dfrac{y}{x}e^{xy}\, dy = 0$

b) $\dfrac{e^x + y^3}{y^2}\, dx + 3x\, dy = 0$

c) $(4x^5 + 9x^4y^2 + 10\dfrac{y}{x})\, dx + (6x^5y + 3x^2y^2 - 5)\, dy = 0$

d) $2\sin(y)\, dx + x\cos(y)\, dy = 0$

Author(s): Matthew Charnley and Jason Nowell

Exercise 177 *Find the integrating factor for the following equations making them into exact equations:*

a) $\frac{1}{y}\,dx + 3y\,dy = 0$

b) $dx - e^{-x-y}\,dy = 0$

c) $\left(\frac{\cos(x)}{y^2} + \frac{1}{y}\right)dx + \frac{x}{y^2}\,dy = 0$

d) $\left(2y + \frac{y^2}{x}\right)dx + (2y + x)\,dy = 0$

Exercise 178 *Suppose you have an equation of the form:* $f(x) + g(y)\frac{dy}{dx} = 0$.

a) *Show it is exact.*

b) *Find the form of the potential function in terms of* f *and* g.

Exercise 179 *Suppose that we have the equation* $f(x)\,dx - dy = 0$.

a) *Is this equation exact?*

b) *Find the general solution using a definite integral.*

Exercise 180 *Find the potential function* $F(x, y)$ *of the exact equation* $\frac{1+xy}{x}\,dx + \left({}^{1}\!/_{y} + x\right)dy = 0$ *in two different ways.*

a) *Integrate* M *in terms of* x *and then differentiate in* y *and set to* N.

b) *Integrate* N *in terms of* y *and then differentiate in* x *and set to* M.

Exercise 181 *A function* $u(x, y)$ *is said to be a* harmonic function *if* $u_{xx} + u_{yy} = 0$.

a) *Show if* u *is harmonic,* $-u_y\,dx + u_x\,dy = 0$ *is an exact equation. So there exists (at least locally) the so-called* harmonic conjugate *function* $v(x, y)$ *such that* $v_x = -u_y$ *and* $v_y = u_x$.

Verify that the following u *are harmonic and find the corresponding harmonic conjugates* v*:*

b) $u = 2xy$

c) $u = e^x \cos y$

d) $u = x^3 - 3xy^2$

Exercise 182

a) *Show that every separable equation* $y' = f(x)g(y)$ *can be written as an exact equation, and verify that it is indeed exact.*

b) *Using this rewrite* $y' = xy$ *as an exact equation, solve it and verify that the solution is the same as it was in* **Normally a reference to a previous example goes here.**.

Modeling with First Order Equations

Stuff about Modeling with First Order Equations

One of the main reasons to study and learn about differential equations, particularly for scientists and engineers, is their application and use in mathematical modeling. Since the derivative of a function represents the rate of change of that quantity, if we can use physical or scientific principles to develop an equation for the rate of change of some quantity in terms of the quantity and time, there's a chance that we can write a differential equation for this quantity and solve it to determine how the quantity will change.

Principles of Mathematical Modeling

The process of mathematical modeling involves three main steps. The first of these is to write the model. This part comes from basic science or engineering principles and involves writing a differential equation that fits the given situation. If we can determine the rate at which a quantity will change based on the surrounding factors, we have a good shot of getting to such an equation. One main principle that can be used to write these equations is the accumuilation equation, which will be discussed in the next subsection.

The second step of this process is to solve the differential equation. This can mean either an analytic solution or a numeric one, and this is where the work of this class comes into play. We are going through a bunch of different techniques for solving differential equations and analyzing the overall behavior of such equations so that we can use them in this way. The end goal is to get an equation or a graph for how the quantity that we made a model for is going to change in time.

The final step of the process is to validate the model by comparing with experimental data. Once we have written the model and solved the corresponding differential equation, we want to make sure that the model works. To do this, we can take a new version of the original scenario, run the model as well as the physical experiment and see how the results compare. If the results are "close" (in whatever sense makes logical sense for the problem), then we have a good model and can keep it. However, if our results differ significantly, then the model we used probably doesn't apply to this problem. We need to go back to step 1 to try to figure out a better model for the physical situation in order to get more accurate results.

Why do we care about mathematical modeling? The biggest thing that it does from an engineering point of view is reduce the need for repeated testing. If we have a mathematical model that works for a given physical system, we can see how the system will be have under slightly different conditions and with different initial conditions without needing to run the physical experiment over and over again. We can do all of this testing on the model, and since we have validated the model, we can assume that the actual results will be similar. This also allows us to change some aspects of the physical situation to try to optimize it, but do so just by modifying the mathematical model, not the physical setup. This can significantly cut down on costs and allow for more optimal system design at the same time.

The Accumulation Equation

The accumulation equation is one of the simplest general mathematical formulations that can be used to develop mathematical models. This equation comes down to the fact that the rate of change of some quantity should be equal to the rate at which it is being added minus the rate at which it is being removed. If we let x be the quantity in question, this can be written as

$$\frac{dx}{dt} = \text{rate in} - \text{rate out}. \tag{14}$$

This may seem fairly simple. However, it shows up in many places in science and engineering. Any mass or energy balance equations are examples of accumulation equations. These types of equations can also be written for the accumulation of momentum, and doing so for fluids gives rise to the Navier-Stokes equations, providing the basis for several fields of engineering. The examples that we see here will be simpler than that, but the idea is still the same.

Example 31. *A tank initially contains 70 gallons of water and 5 lbs of salt. A solution with salt concentration 0.2 lbs per gallon flows into the tank at a rate of* 3 *gal/min. The tank is well stirred, and water is removed from the tank at a rate of* 3 *gal/min. Find the amount of salt in the tank at any time t? What happens as $t \to \infty$? Does this make sense?*

Learning outcomes: Write a first-order differential equation to model a physical situation Interpret the solution to a differential equation in the context of a physical problem.

Author(s): Matthew Charnley and Jason Nowell

Solution: To solve this problem, we use the accumulation equation (14) on the amount of salt in the tank. In order to compute with this, we recognize that in terms of mass of salt moving into the tank

$$\text{rate in} = \text{flow in} \times \text{concentration in}$$

and similarly for the mass of salt leaving the tank.

If we let x represent the amount of salt in the tank at any time t (which is the goal of the problem), we can write a differential equation for this using the accumulation equation (14). This gives us that

$$\frac{dx}{dt} = \text{rate in} - \text{rate out} = \text{flow in} \times \text{concentration in} - \text{flow out} \times \text{concentration out}$$

For this problem, we have that

$$\begin{aligned} \text{flow in} &= 3, \\ \text{concentration in} &= 0.2, \\ \text{flow out} &= 3, \\ \text{concentration out} &= \frac{x}{\text{volume}} = \frac{x}{70}. \end{aligned}$$

The last of these lines comes from the fact that the tank is "well stirred" or "well-mixed." This implies that the concentration of salt in the water leaving the tank is the same as the concentration in the tank, which we can compute as $\frac{x}{\text{volume}}$. In this case, since the flow rate in and out are both 3 gal/min, the volume of water in the tank is fixed at 70 gallons, so we can put this in the equation.

Therefore, our equation becomes

$$\frac{dx}{dt} = (3 \times 0.2) - \left(3 \times \frac{x}{70}\right).$$

We can rewrite this equation as

$$\frac{dx}{dt} + \frac{3}{70}x = 0.6$$

which we recognize as a first order linear equation. We can then solve this using the method of integrating factors. Our factor $r(t)$ is

$$r(t) = e^{\int p(t)\,dt} = e^{\int \frac{3}{70}\,dt} = e^{\frac{3}{70}t},$$

which we can multiply on both sides of the equation to obtain

$$e^{\frac{3}{70}t}\frac{dx}{dt} + e^{\frac{3}{70}t}\frac{3}{70}x = 0.6e^{\frac{3}{70}t}.$$

The left side of this is a product rule derivative, so we can integrate both sides to obtain

$$e^{\frac{3}{70}t}x = 0.6\frac{70}{3}e^{\frac{3}{70}t} + C.$$

We can then isolate x to get our general solution as

$$x = 14 + Ce^{-\frac{3}{70}t}.$$

Our initial condition tells us that $x(0) = 5$. Plugging this in gives that

$$5 = x(0) = 14 + C \quad \Rightarrow \quad C = -9,$$

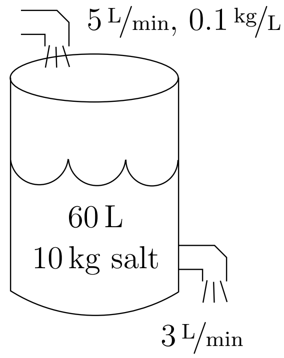
so the solution to the initial value problem, and thus our calculation for the amount of salt in the tank at any time t, is

$$x(t) = 14 - 9e^{-\frac{3}{70}t}.$$

As $t \to \infty$, we see that the exponential term goes to zero. This leaves us with 14 lbs of salt in the tank after a long time. This makes some sense because this would give us a concentration of $\frac{14}{70} = 0.2$ lb/gal, and that was exactly the concentration of the in-flow stream. It makes sense that after a long time of mixing and removing water from the tank, the concentration of the tank would match that of the incoming stream.

The same principle works for other types of examples, including those where the volume of the tank is not constant in time.

Example 32. *A 100 liter tank contains 10 kilograms of salt dissolved in 60 liters of water. Solution of water and salt (brine) with concentration of 0.1 kilograms per liter is flowing in at the rate of 5 liters a minute. The solution in the tank is well stirred and flows out at a rate of 3 liters a minute. How much salt is in the tank when the tank is full?*



Solution: We can again use the accumulation equation to write

$$\frac{dx}{dt} = (\text{flow in} \times \text{concentration in}) - (\text{flow out} \times \text{concentration out}).$$

In this example, we have

$$\begin{aligned} \text{flow in} &= 5, \\ \text{concentration in} &= 0.1, \\ \text{flow out} &= 3, \\ \text{concentration out} &= \frac{x}{\text{volume}} = \frac{x}{60 + (5-3)t}. \end{aligned}$$

Our equation is, therefore,

$$\frac{dx}{dt} = (5 \times 0.1) - \left(3 \frac{x}{60+2t} \right).$$

Or in the form (8)

$$\frac{dx}{dt} + \frac{3}{60+2t} x = 0.5.$$

Let us solve. The integrating factor is

$$r(t) = \exp \left(\int \frac{3}{60+2t} dt \right) = \exp \left(\frac{3}{2} \ln(60+2t) \right) = (60+2t)^{3/2}.$$

We multiply both sides of the equation to get

$$\begin{aligned} (60+2t)^{3/2} \frac{dx}{dt} + (60+2t)^{3/2} \frac{3}{60+2t} x &= 0.5(60+2t)^{3/2}, \\ \frac{d}{dt} \left[(60+2t)^{3/2} x \right] &= 0.5(60+2t)^{3/2}, \\ (60+2t)^{3/2} x &= \int 0.5(60+2t)^{3/2} dt + C, \\ x &= (60+2t)^{-3/2} \int \frac{(60+2t)^{3/2}}{2} dt + C(60+2t)^{-3/2}, \\ x &= (60+2t)^{-3/2} \frac{1}{10} (60+2t)^{5/2} + C(60+2t)^{-3/2}, \\ x &= \frac{60+2t}{10} + C(60+2t)^{-3/2}. \end{aligned}$$

We need to find C. We know that at $t = 0$, $x = 10$. So

$$10 = x(0) = \frac{60}{10} + C(60)^{-3/2} = 6 + C(60)^{-3/2},$$

or

$$C = 4(60^{3/2}) \approx 1859.03.$$

We are interested in x when the tank is full. The tank is full when $60 + 2t = 100$, or when $t = 20$. So

$$\begin{aligned} x(20) &= \frac{60+40}{10} + C(60+40)^{-3/2} \\ &\approx 10 + 1859.03(100)^{-3/2} \approx 11.86. \end{aligned}$$

See **Normally a reference to a previous figure goes here.** for the graph of x over t.

The concentration when the tank is full is approximately 0.1186 kg/liter, and we started with $1/6$ or 0.167 kg/liter.

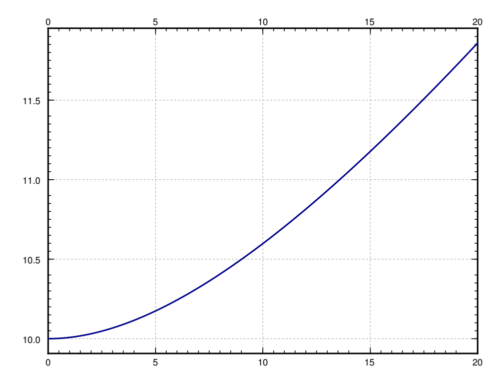


Figure 33: Graph of the solution x kilograms of salt in the tank at time t.

For the previous example, we obtained the solution

$$x(t) = \frac{60+2t}{10} + 1859.03(60+2t)^{-3/2},$$

which is valid and well defined for all positive values of t (it has an issue at $t = -30$, but we aren't concerned about that here). However, as a differential equation that represents a physical situation, it is not valid for all positive values of t. The issue here is that the tank is full at $t = 20$. Therefore, beyond this point, while the function still exists, it is not a valid model for this physical system. Once the tank fills, you can't keep adding and removing water at the same rates that you have been up until this point, because something is going to break with the system. The same goes for if you are removing water from the tank at a faster rate than you are adding it, because then the tank will empty at some point in time and beyond that, the model equation no longer represents the system.

The same ideas apply to problems involving interest compounded continuously. For an interest rate of r, the "rate in," or the rate at which the money in the account is increasing, is rP where P is the amount of money in the account. Taking this along with other factors that may affect the balance of the account allows us to write a differential equation, which we can solve to determine what the balance will be over time.

Example 33. *A bank account with an interest rate of 6% per year, compounded continuously, starts with a balance of \$30000. The owner of the account withdraws \$50 from the account each month. Find and solve a differential equation for the account balance over time. What is the largest amount that the owner could withdraw each month without the account eventually reaching \$0?*

Solution: We will use the function $P(t)$ to model the balance of the account over time, where t is in *years*. Since the owner withdraws \$50 per month, this means that they withdraw \$600 over the course of the year. This means that the differential equation we want is

$$\frac{dP}{dt} = 0.06P - 600 \qquad P(0) = 30000.$$

We can solve this equation by the integrating factor method.

$$\begin{aligned} P' - 0.06P &= -600 \\ (e^{-0.06t}P)' &= -600e^{-0.06t} \\ e^{-0.06t}P &= 10000e^{-0.06t} + C \\ P &= 10000 + Ce^{0.06t} \end{aligned}$$

For $P(0) = 30000$, we need to take $C = 20000$. Thus, the solution to the initial value problem is

$$P(t) = 10000 + 20000e^{0.06t}.$$

Since the coefficient in front of $e^{0.06t}$ is positive, this means that the account balance here will grow in time.

For the second part, we need to adjust the withdrawal amount to see how the solution changes. If we let K be the monthly withdrawal amount, then we have the differential equation

$$\frac{dP}{dt} = 0.06P - 12K \qquad P(0) = 30000.$$

The same solution method gives us

$$P(t) = \frac{12K}{0.06} + Ce^{0.06t}.$$

If $C < 0$, then the account balance will eventually go to zero. Therefore, we need $C \geq 0$, and since $P(0) = 30000$, we have that

$$30000 = \frac{12K}{0.06} + C \quad \text{or} \quad C = 30000 - \frac{12K}{0.06}.$$

For this to be equal to zero, we need

$$\frac{12K}{0.06} = 30000 \qquad K = 150.$$

Thus, the owner can withdraw \$150 per month and keep the account balance positive.

To end this section, we will analyze the example that was presented at the very beginning of the book.

Example 34. *An object falling through the air has its velocity affected by two factors: gravity and a drag force. The velocity downward is increased at a rate of* $9.8\ m/s^2$ *due to gravity, and it is decreased by a rate equation to* 0.3 *times the current velocity of the object. If the ball is initially thrown downwards at a speed of* $2\ m/s$*, what will the velocity be 10 seconds later?*

Solution: As described in that first section, we know that the differential equation that we can write for this situation is

$$\frac{dv}{dt} = 9.8 - 0.3v$$

and that the initial condition for the velocity if $v(0) = 2$. Since we have gravity as a positive 9.8, this means that the downward direction is positive, so the object being thrown downward at $2\ m/s$ means that it is positive. We then need to solve this initial value problem, which we can do using first order linear methods. The equation can be written as

$$v' + 0.3v = 9.8$$

which has integrating factor $e^{0.3t}$. After multiplying this to both sides and integrating, we get that

$$e^{0.3t}v = \frac{9.8}{0.3}e^{0.3t} + C$$

or that

$$v(t) = \frac{9.8}{0.3} + Ce^{-0.3t}.$$

Using the initial condition, we get that

$$v(0) = \frac{98}{3} + C = 2$$

so that $C = -\frac{92}{3}$ and the solution to the initial value problem is

$$v(t) = \frac{98}{3} - \frac{92}{3}e^{-0.3t}.$$

Then, to determine the velocity at $t = 10$, we can plug 10 into this formula to get that

$$v(10) = \frac{98}{3} - \frac{92}{3}e^{-3} \approx 31.14\ m/s.$$

All of these examples are based around the same idea of the accumulation equation. We need to determine the quantity that is changing as well as all of the factors that cause it to increase and decrease. These get combined into a differential equation which we can solve in order to analyze the situation and answer whatever questions you want about that physical problem. Keeping these ideas in mind will help you approach a wide variety of problems both in this class as well as future applications in engineering classes and beyond.

`firstOrder/practice/modeling-practice1.tex`

Practice for Modeling

Why?

Exercise 183 *Suppose there are two lakes located on a stream. Clean water flows into the first lake, then the water from the first lake flows into the second lake, and then water from the second lake flows further downstream. The in and out flow from each lake is 500 liters per hour. The first lake contains 100 thousand liters of water and the second lake contains 200 thousand liters of water. A truck with 500 kg of toxic substance crashes into the first lake. Assume that the water is being continually mixed perfectly by the stream.*

a) *Find the concentration of toxic substance as a function of time in both lakes.*

b) *When will the concentration in the first lake be below 0.001 kg per liter?*

c) *When will the concentration in the second lake be maximal?*

Exercise 184 *Newton's law of cooling states that* $\frac{dx}{dt} = -k(x - A)$ *where* x *is the temperature,* t *is time,* A *is the ambient temperature, and* $k > 0$ *is a constant. Suppose that* $A = A_0 \cos(\omega t)$ *for some constants* A_0 *and* ω*. That is, the ambient temperature oscillates (for example night and day temperatures).*

a) *Find the general solution.*

b) *In the long term, will the initial conditions make much of a difference? Why or why not?*

Exercise 185 *Initially 5 grams of salt are dissolved in 20 liters of water. Brine with concentration of salt 2 grams of salt per liter is added at a rate of 3 liters per minute. The tank is mixed well and is drained at 3 liters per minute. How long does the process have to continue until there are 20 grams of salt in the tank?*

Exercise 186 *Initially a tank contains 10 liters of pure water. Brine of unknown (but constant) concentration of salt is flowing in at 1 liter per minute. The water is mixed well and drained at 1 liter per minute. In 20 minutes there are 15 grams of salt in the tank. What is the concentration of salt in the incoming brine?*

Exercise 187 *Suppose a water tank is being pumped out at* $3\,\mathrm{L/min}$*. The water tank starts at* $10\,\mathrm{L}$ *of clean water. Water with toxic substance is flowing into the tank at* $2\,\mathrm{L/min}$*, with concentration* $20t\,\mathrm{g/L}$ *at time* t*. When the tank is half empty, how many grams of toxic substance are in the tank (assuming perfect mixing)?*

Exercise 188 *A 300 gallon well-mixed water tank initially starts with 200 gallons of water and 15 lbs of salt. One stream with salt concentration one pound per gallon flows into the tank at a rate of 3 gallons per minute and water is removed from the well-mixed tank at a rate of 2 gallons per minute.*

a) *Write and solve an initial value problem for the volume of water in the tank at any time* t*.*

b) *Set up an initial value problem for the amount of salt in the tank at any time* t*. You do not need to solve it (yet), but should make sure to state it fully.*

Author(s): Matthew Charnley and Jason Nowell

c) *Is the solution to this initial value problem a valid representation of the physical model for all times $t > 0$? If so, use the information in the equation to determine the long-time behavior of the solution. If not, explain why, determine the time when the representation breaks down, and what happens at that point in time.*

d) *Solve the initial value problem above and compare this to your answer to the previous part.*

Exercise 189 *A 500 gallon well-mixed water tank initially starts with 300 gallons of water and 200 lbs of salt. One stream with salt concentration of 0.5* $^{\text{lb}}/_{\text{gal}}$ *flows into the tank at a rate of 5* $^{\text{gal}}/_{\text{min}}$ *and water is removed from the well-mixed tank at a rate of 7* $^{\text{gal}}/_{\text{min}}$.

a) *Write and solve an initial value problem for the volume of water in the tank at any time t.*

b) *Set up an initial value problem for the amount of salt in the tank at any time t. You do not need to solve it (yet), but should make sure to state it fully.*

c) *Is the solution to this initial value problem a valid representation of the physical model for all times $t > 0$? If so, use the information in the equation to determine the long-time behavior of the solution. If not, explain why, determine the time when the representation breaks down, and what happens at that point in time.*

d) *Solve the initial value problem above and compare this to your answer to the previous part.*

Exercise 190 *A 200 gallon well-mixed water tank initially starts with 150 gallons of water and 50 lbs of salt. One stream with salt concentration of 0.2* $^{\text{lb}}/_{\text{gal}}$ *flows into the tank at a rate of 4* $^{\text{gal}}/_{\text{min}}$ *and water is removed from the well-mixed tank at a rate of 4* $^{\text{gal}}/_{\text{min}}$.

a) *Write and solve an initial value problem for the volume of water in the tank at any time t.*

b) *Set up an initial value problem for the amount of salt in the tank at any time t. You do not need to solve it (yet), but should make sure to state it fully.*

c) *Is the solution to this initial value problem a valid representation of the physical model for all times $t > 0$? If so, use the information in the equation to determine the long-time behavior of the solution. If not, explain why, determine the time when the representation breaks down, and what happens at that point in time.*

d) *Solve the initial value problem above and compare this to your answer to the previous part.*

Exercise 191 *Suppose we have bacteria on a plate and suppose that we are slowly adding a toxic substance such that the rate of growth is slowing down. That is, suppose that* $\frac{dP}{dt} = (2 - 0.1\,t)P$. *If* $P(0) = 1000$, *find the population at* $t = 5$.

Exercise 192 *A cylindrical water tank has water flowing in at I cubic meters per second. Let A be the area of the cross section of the tank in meters. Suppose water is flowing from the bottom of the tank at a rate proportional to the height of the water level. Set up the differential equation for h, the height of the water, introducing and naming constants that you need. You should also give the units for your constants.*

Exercise 193 *An object in free fall has a velocity that increases at a rate of 32* ft/s^2. *Due to drag, the velocity decreases at a rate of 0.1 times the velocity of the object squared, when written in feet per second.*

a) *Write a differential equation to model the velocity of this object over time.*

b) *This equation is autonomous, so draw a phase diagram for this equation and classify all critical points.*

c) *What will happen to the velocity if the object is dropped at $t = 0$? What about if the object is thrown downwards at a rate of $10 ft/s$?*

Exercise 194 *The number of people in a town that support a given measure decays at a constant rate of 10 people per day. However, the support for the measure can be increased by individuals discussing the issue. This results in an increase of the support at a rate of $ay(1000 - y)$ people per day, where y is the number of people who support the measure, and a is a constant depending on the way in which the issue is being discussed. Write a differential equation to model this situation, and determine the amount of people who will support the measure long-term if a is set to 10^{-4}.*

Exercise 195 *Newton's Law of Procrastination states that the rate at which one accomplishes a chore is* proportional to the amount of the chore not yet done. *Unbeknownst to Newton, this applies to robots too. A Roomba is attempting to vacuum a house measuring 1000 square feet. When none of the house is clean, the roomba can clean 200 square feet per hour. What makes this problem fun is that there is also a dog. It's whatever kind of dog you like, take your pick. The dog dirties the house at a constant rate of 50 square feet per hour.*

a) *Assume that none of the house is clean at $t = 0$. Write a DE for the number of square feet that are clean as a function of time, and solve for that quantity.*

b) *How long will it take before the house is half clean? Will it ever be entirely clean? (Explain briefly.)*

Exercise 196 *A student has a loan for \$50000 with 5% interest. The student makes \$300 payments on the loan each month. The rate here is an annual rate, compounded continuously, and the differential equation you write should be in years.*

a) *With this setup, how long does it take the student to pay off the loan? How much money does the student pay over this period of time?*

b) *What is the minimal amount the student should pay each month if they want to pay off the loan within 5 years? How much does the student pay over this period?*

Exercise 197 *A factory pumps 6 tons of sludge per day into a nearby pond. The pond initially contains 100,000 gallons of water, and no sludge. Each day, 3,000 gallons of rain water falls into the pond, and 1,000 gallons per day leave the pond via a river. Assume, for no good reason, that the water leaving the pond has the same concentration of sludge as the pond as a whole. How much sludge will there be in the pond after 150 days?*

Exercise 198 *In this exercise, we compare two different young people and their investment strategies. Both of these people are investing in an account with 7.5% annual rate of return. Person 1 invests \$50 a month starting at age 20, and Person 2 invests \$100 per month starting at age 30. Write differential equations to model each of these account balances over time, and compute the amount of money in each account at age 50. Who has more money in the account? Who has invested more money? What would person 2 have to invest each month in order for the two balances to be equal at age 50?*

Exercise 199 *Radioactive decay follows similar rules to interest, where a certain portion of the material decays over time, resulting in an equation of the form*

$$\frac{dy}{dt} = -ky$$

for some constant k. The half-life *of a material is the amount of time that it takes for half of the material to have decayed away. Assume that the* half-life *of a given substance is T minutes. Find a formula for k, the coefficient in the decay equation, in terms of T.*

Modeling and Parameter Estimation

We discuss Modeling and Parameter Estimation

One of the most common ways that the mathematical modeling structure can be used to analyze physical problems is the idea of parameter estimation. The situation is that we have physical principles that give rise to a differential equation that defines how a physical system should behave, but there are one or more constants in the problem that we do not know. Two simpler examples of this are Newton's Law of Cooling

$$\frac{dT}{dt} = -k(T - T_s)$$

which models the temperature of an object in an environment of temperature T_s over time, and velocity affected by drag

$$\frac{dv}{dt} = 9.8 - \alpha v^2$$

modeling the velocity of a falling object where the drag force is proportional to the square of the velocity. In both of these cases, the models are well established, but for a given object, we likely do not know the k or α values in the problem. These are these *parameters* of the problem, and would be determined by the shape and structure of the objects, the material that it is made of, and many other factors, so it could be hard to figure out what they are in advance. How can we find these values? We can use data from the actual physical problem to try to estimate these parameters.

The easier version of this is to use a single value at a later time to calculate the constant.

Example 35. *An object that obeys Newton's Law of Cooling is placed in an environment at a constant temperature of 20^o C. The object starts at 50^o C, and after 10 minutes, it has reached a temperature of 40^o C. Find a function for the temperature as a function of time.*

Solution: Based on Newton's Law of Cooling, we know that the temperature satisfies the differential equation

$$\frac{dT}{dt} = -k(T - T_s) = -k(T - 20)$$

with initial condition $T(0) = 50$, but we do not know the value of k. In order to work this out, we should solve the differential equation with unknown constant k, then figure out which value of k gives us the appropriate temperature after 10 minutes. This is a first order linear equation, which can be rewritten as

$$T' + kT = 20k.$$

The integrating factor we need is e^{kt}, which turns the equation into

$$(e^{kt}T)' = 20ke^{kt}.$$

Integrating both sides and solving for T gives

$$T(t) = 20 + Ce^{-kt}.$$

To satisfy the initial condition, we need that $T(0) = 50$, or $C = 30$. Thus, our solution, still with an unknown constant k, is

$$T(t) = 20 + 30e^{-kt}.$$

To determine the value of k, we need to utilize the other given piece of information: that $T(10) = 40$. Plugging this in gives that

$$40 = 20 + 30e^{-10k}$$

which we can solve for k using logarithms. This will give that

$$\frac{2}{3} = e^{-10k} \qquad \Rightarrow \qquad k = -\frac{1}{10}\ln\frac{2}{3}.$$

Learning outcomes: Use parameter estimation to approximate physical parameters from data.
Author(s): Matthew Charnley and Jason Nowell

Finally, we can plug that constant into our equation to get the solution for the temperature at any time value,

$$T(t) = 20 + 30e^{\frac{t}{10}\ln\frac{2}{3}}.$$

This method works great if we have the exact measurement from the object at one point in time. However, if the measurements at multiple points in time are known, and if the data is not likely to be exact, then a different method is more applicable. The idea is that we want to minimize the "error" between our predicted result and the physical data that we gather. The method used to minimize the error is the "Least Squared Error" method.

Assume that we want to do this for the drag coefficient problem,

$$\frac{dv}{dt} = 9.8 - \alpha v^2$$

where we do not know, and want to estimate, the value of α. For this method, the data that we gather is a set of velocity values $v_1, v_2, ..., v_n$ that are obtained at times $t_1, t_2, ..., t_n$. For any given value of α, we can solve, either numerically or analytically, the solution v_α to the given differential equation with that value of α. From this solution, we can compute $v_\alpha(t_1), v_\alpha(t_2), ..., v_\alpha(t_n)$, the value of this solution at each of the times that we gathered data originally. Now, we want to compute the error that we made in choosing this parameter α. This is computed by

$$E(\alpha) = (v_1 - v_\alpha(t_1))^2 + (v_2 - v_\alpha(t_2))^2 + \cdots + (v_n - v_\alpha(t_n))^2$$

which is the sum of the squares of the differences between the gathered data and the predicted solution. In order to find the best possible value of α, we want to minimze this error by choosing different values of α

$$E_{min} = \min_\alpha E(\alpha) = \min_\alpha \sum_{i=1}^{n} (v_i - v_\alpha(t_i))^2$$

and whatever value of α gives us this minimum is the optimal choice for that parameter.

The function that we want to minimize here is usually a very complicated function, and we may not even be able to solve the differential equation analytically for any α. Thus, computers are used most often here to solve these types of problems.

Example 36. *An object is falling under the force of gravity, and has a drag component that is proportional to the square of the velocity. Data is gathered on the falling object, and the velocity at a variety of times are given in* ***Normally a reference to a previous table goes here.****.*

t (s)	**v** (m/s)
0	0
0.1	0.9797
0.3	2.8625
0.5	4.4750
0.8	6.3828
0.9	6.8360
1.0	7.0334
1.5	8.1612

Table 3: Data for estimating drag coefficient using least squared errors.

Use this data to estimate the coefficient of proportionality on the drag term in the equation

$$\frac{dv}{dt} = 9.8 - \alpha v^2.$$

Solution: To solve this problem, we will use the least squared error method implemented in MATLAB. The code we need for this is the following, which makes use of the Optimization Toolbox.

Code

```
    global tVals
    global vVals

```

```
tVals = [0, 0.1, 0.3, 0.5, 0.8, 0.9, 1.0, 1.5];
vVals = [0,0.9797,2.8625,4.4750,6.3828,6.8360,7.0334,8.1612];

[aVal, errVal] = fminbnd(@(a) EstSqError(a), 0, 4)
```

This bit of code inputs the necessary values and uses the `fminbnd` function to find the minimum of the error function on a defined interval. These problems need to be done on a bounded interval, but in most physical situations there is some reasonable window for where the parameter could be. The rest of the code is the definition of the `EstSqError` function.

Code

```
function err = EstSqError(al)

global tVals
global vVals

fun = @(t,v) 9.8 - al.*v.^2;
sol = ode45(fun, [0,3], 0);
vTest = deval(sol, tVals);

err = sum((vVals - vTest).^2)
end
```

The main point of this code is that it takes in a value of α, over which we are trying to minimize, it numerically solves the differential equation with that value of α over a desired range of values, and then compares the inputted `vVals` with the generated `vTest` array, computing the sum of squared errors, and returning the error value.

Running this code results in an α value of 0.1256, with an error of 0.0345. That means that, based on this data, the best approximation to α is 0.1256.

Note that in the above example, the total error was not zero, and doesn't actually match the coefficient used to generate the data, which was 0.123. This is because noise was added to the data before trying to compute the drag coefficient. In a real world problem, noise would not be added, but a similar effect would arise from slightly inaccurate measurements or round-off errors in the data. While we may not have found the constant exactly, we got really close to it, and could use this as a starting point for further experiments and data validation.

firstOrder/practice/modelingParameters-practice1.tex

Practice for Modeling Parameters

Why?

Exercise 200 *Bob is getting coffee from a restaurant and knows that the temperature of the coffee will follow Newton's Law of Cooling, which says that*

$$\frac{dT}{dt} = k(T_0 - T)$$

where T_0 is the ambient temperature and k is a constant depending on the object and geometry. His car is held at a constant 20° C, and when he receives the coffee, he measures the temperature to be 90° C. Two minutes later, the temperature is 81°C.

a) *Use these two points of data to determine the value of k for this coffee.*

b) *Bob only wants to drink his coffee once it reaches 65° C. How long does he have to wait for this to happen?*

c) *If the coffee is too cold for Bob's taste once it reaches 35° C, how long is the acceptable window for Bob to drink his coffee?*

Exercise 201 *Assume that a falling object has a velocity (in meters per second) that obeys the differential equation*

$$\frac{dv}{dt} = 9.8 - \alpha v$$

where α represents the drag coefficient of the object.

a) *Solve this differential equation with initial condition $v(0) = 0$ to get a solution that depends on α.*

b) *Assume that you drop an object from a height of 10 meters and it hits the ground after 3 seconds. What is the value of α here? (Note: You solved for $v(t)$ in the previous part, and now you need to get to position. What does that require?)*

c) *Assume that a second object hits the ground in 6 seconds. How does this change the value of α?*

Exercise 202 *A restaurant is trying to analyze the to-go coffee cups that it uses in order to best serve their customers. They assume that the coffee follows Newton's Law of Cooling and place it in a room with ambient temperature 22° C. They record the following data for the temperature of the coffee as a function of time.*

t (min)	**T** (° C)
0	80
0.5	74
1.1	67
1.7	61
2.3	56

a) *Use code to determine the best-fit value of k for this data.*

b) *The restaurant determines that to avoid any potential legal issues, the coffee can be no warmer than 60 °C when it is served. If the coffee comes out of the machine at 90 °C, how long do they have to wait before they can serve the coffee?*

Author(s): Matthew Charnley and Jason Nowell

Exercise 203 *Assume that an object falling has a velocity that follows the differential equation*

$$\frac{dv}{dt} = 9.8 - \alpha v^2$$

where the velocity is in $^{m}/_{s}$ *and* α *represents the drag coefficient. During a physics experiment, a student measures data for the velocity of a falling object over time given in the table below.*

Use this data (and code) to estimate the value of α *for this object.*

t (s)	**v** (m/s)
0	0
0.1	1.0
0.2	1.9
0.4	3.6
0.6	5.2
0.9	6.8
1.1	7.4
1.3	7.9
1.5	8.2
1.8	8.5
2.1	8.8

Table 4: Data for Exercise .

t (d)	**P** (thousands)
0	50
7	60
14	70
28	97
37	117
50	148
78	220
100	268

Table 5: Data for Exercise .

Exercise 204 *Assume that a species of fish in a lake has a population that is modeled by the differential equation*

$$\frac{dP}{dt} = \frac{1}{100} rP(K - P) - \alpha$$

where r*,* K*, and* α *are parameters,* r *representing the growth rate,* K *the carrying capacity, and* α *the harvesting rate, and the population* P *is in thousands., with* t *given in years. From previous studies, you know that the best value of* r *is* 3.12*. After studying the population over a period of time, you get the data given above.*

a) *Your friend tells you that in a previous study, he found that the value of* K *for this particular lake is* 450*. Use code to determine the best value of* α *for this situation. Note that the equation expects* t *in years, but the data is given in days. Search for* α *in the range* $(0, 400)$*.*

b) *That answer doesn't look great. Plot the solution with these parameters along with the data and compare them.*

c) *The fit does not look great, so maybe your friend's value was not quite right. Run code to find best values for* K *and* α *simultaneously. Use the range* $(0, 400)$ *for both* α *and* K*.*

Substitution

Stuff about Substitution

The equation

$$y' = (x - y + 1)^2$$

is neither separable nor linear. What can we do? One technique that worked for helping us in evaluating integrals was substitution, or change of variables. For example, in order to evaluate the integral

$$\int 2x(x^2+4)^5 \; dx$$

we can do so by defining $u = x^2 + 4$ so that $du = 2x \; dx$, and then evaluate the integral as

$$\int u^5 \; du = \frac{u^6}{6} + C = \frac{(x^2+4)^6}{6} + C$$

after we have plugged our original function back in.

We can try to do the same thing here, and be careful with how we set things up. Our general strategy will be to pick a new dependent variable, find a differential equation that this new variable solves (which will come from the old equation), solve that equation, then convert back to the original variable. We will take v as our new dependent variable here, which is as function $v(x)$. Let us try

$$v = x - y + 1,$$

which is the "inside" function (it's inside the square) of our example. In order to get to a differential equation that v satisfies, we need to figure out y' in terms of v', v and x. We differentiate (in x) to obtain $v' = 1 - y'$. So $y' = 1 - v'$. We plug this into the equation to get

$$1 - v' = v^2.$$

In other words, $v' = 1 - v^2$. Such an equation we know how to solve by separating variables:

$$\frac{1}{1-v^2} \, dv = dx.$$

So

$$\frac{1}{2} \ln \left| \frac{v+1}{v-1} \right| = x + C, \qquad \text{or} \qquad \left| \frac{v+1}{v-1} \right| = e^{2x+2C}, \qquad \text{or} \qquad \frac{v+1}{v-1} = De^{2x},$$

for some constant D. Note that $v = 1$ and $v = -1$ are also solutions; they are the *singular solutions* in this problem. (This solution method requires partial fractions; for a review of that topic, see § **??**.)

Now we need to "unsubstitute" to obtain

$$\frac{x-y+2}{x-y} = De^{2x},$$

and also the two solutions $x - y + 1 = 1$ or $y = x$, and $x - y + 1 = -1$ or $y = x + 2$. We solve the first equation for y.

$$\begin{aligned} x - y + 2 &= (x-y)De^{2x}, \\ x - y + 2 &= Dxe^{2x} - yDe^{2x}, \\ -y + yDe^{2x} &= Dxe^{2x} - x - 2, \\ y\left(-1 + De^{2x}\right) &= Dxe^{2x} - x - 2, \\ y &= \frac{Dxe^{2x} - x - 2}{De^{2x} - 1}. \end{aligned}$$

Note that $D = 0$ gives $y = x + 2$, but no value of D gives the solution $y = x$.

Substitution in differential equations is applied in much the same way that it is applied in calculus. You guess. Several different substitutions might work. There are some general patterns to look for. We summarize a few of these in a table.

Learning outcomes: Use substitution to solve more complicated first order equations Use a Bernoulli substitution to solve appropriate first order equations Use a homogeneity transformation to solve appropriate first order equations.

Author(s): Matthew Charnley and Jason Nowell

When you see	Try substituting
yy'	$v = y^2$
y^2y'	$v = y^3$
$(\cos y)y'$	$v = \sin y$
$(\sin y)y'$	$v = \cos y$
$y'e^y$	$v = e^y$

Usually you try to substitute in the "most complicated" part of the equation with the hopes of simplifying it. The table above is just a rule of thumb. You might have to modify your guesses. If a substitution does not work (it does not make the equation any simpler), try a different one.

Bernoulli equations

There are some forms of equations where there is a general rule for substitution that always works. One such example is the so-called *Bernoulli equation*[1]:

$$y' + p(x)y = q(x)y^n.$$

This equation looks a lot like a linear equation except for the y^n. If $n = 0$ or $n = 1$, then the equation is linear and we can solve it. Otherwise, the substitution $v = y^{1-n}$ transforms the Bernoulli equation into a linear equation. Note that n need not be an integer.

Example 37. *Find the general solution of*

$$y' - \frac{4}{3x}y = -\frac{2}{3}y^4$$

Solution: This equation fits the Bernoulli equation structure with $p(x) = -\frac{4}{3x}$ and $q(x) = -\frac{2}{3}$. Since there is a y^4 on the right-hand side, we take $n = 4$ and make the substitution $v = y^{1-4} = y^{-3}$. With this, we see that

$$v' = -3y^{-4}y'$$

or $y' = -{}^1\!/_3 y^4 v'$. Plugging this into the equation gives

$$\begin{aligned} -\frac{1}{3}y^4v' - \frac{4}{3x}y &= -\frac{2}{3}y^4 \\ -\frac{1}{3}v' - \frac{4}{3x}y^{-3} &= -\frac{2}{3} \\ v' + \frac{4}{x}v &= 2 \end{aligned}$$

This last equation is now a first order linear equation, so we can solve it. The integrating factor we are looking for is

$$\mu(x) = e^{\int p(x)\ dx} = e^{\int \frac{4}{x}\ dx} = e^{4\ln x} = x^4,$$

which results in the euation

$$x^4v' + 4x^3v = 2x^4.$$

The left-hand side is $(x^4v)'$, so we can integrate both sides to get

$$x^4v = \frac{2}{5}x^5 + C,$$

or, solving for v,

$$v(x) = \frac{2}{5}x + \frac{C}{x^4}.$$

However, our original equation was for y, not v. Using the fact that $v = y^{-3}$, we can solve for y as $y = v^{-1/3}$, giving

$$y(x) = \left(\frac{2}{5}x + \frac{C}{x^4}\right)^{-1/3} = \frac{1}{\sqrt[3]{\frac{2}{5}x + \frac{C}{x^4}}}$$

[1]There are several things called Bernoulli equations, this is just one of them. The Bernoullis were a prominent Swiss family of mathematicians. These particular equations are named for Jacob Bernoulli (1654–1705).

as the general solution to this equation.

Even if we need to use integrals to write out the solution to these Bernoulli equations, we can still use the substitution method and solve back out for the desired solution at the end.

Example 38. *Solve*

$$xy' + y(x+1) + xy^5 = 0, \qquad y(1) = 1.$$

Solution: First, the equation is Bernoulli ($p(x) = (x+1)/x$ and $q(x) = -1$). We substitute

$$v = y^{1-5} = y^{-4}, \qquad v' = -4y^{-5}y'.$$

In other words, $(-1/4)\, y^5 v' = y'$. So

$$\begin{aligned} xy' + y(x+1) + xy^5 &= 0, \\ \frac{-xy^5}{4}v' + y(x+1) + xy^5 &= 0, \\ \frac{-x}{4}v' + y^{-4}(x+1) + x &= 0, \\ \frac{-x}{4}v' + v(x+1) + x &= 0, \end{aligned}$$

and finally

$$v' - \frac{4(x+1)}{x}v = 4.$$

The equation is now linear. We can use the integrating factor method. In particular, we use formula (9). Let us assume that $x > 0$ so $|x| = x$. This assumption is OK, as our initial condition is $x = 1$. Let us compute the integrating factor. Here $p(s)$ from formula (9) is $\dfrac{-4(s+1)}{s}$.

$$\begin{aligned} e^{\int_1^x p(s)\,ds} &= \exp\left(\int_1^x \frac{-4(s+1)}{s}\,ds\right) = e^{-4x-4\ln(x)+4} = e^{-4x+4}x^{-4} = \frac{e^{-4x+4}}{x^4}, \\ e^{-\int_1^x p(s)\,ds} &= e^{4x+4\ln(x)-4} = e^{4x-4}x^4. \end{aligned}$$

We now plug in to (9)

$$\begin{aligned} v(x) &= e^{-\int_1^x p(s)\,ds}\left(\int_1^x e^{\int_1^t p(s)\,ds} 4\,dt + 1\right) \\ &= e^{4x-4}x^4\left(\int_1^x 4\frac{e^{-4t+4}}{t^4}\,dt + 1\right). \end{aligned}$$

The integral in this expression is not possible to find in closed form. As we said before, it is perfectly fine to have a definite integral in our solution. Now "unsubstitute"

$$\begin{aligned} y^{-4} &= e^{4x-4}x^4\left(4\int_1^x \frac{e^{-4t+4}}{t^4}\,dt + 1\right), \\ y &= \frac{e^{-x+1}}{x\left(4\int_1^x \frac{e^{-4t+4}}{t^4}\,dt + 1\right)^{1/4}}. \end{aligned}$$

Homogeneous equations

Another type of equations we can solve by substitution are the so-called *homogeneous equations.* Note that this is *not* the same as a homogeneous linear equation. These equations do not have to be linear, and are solved in a very different way. Suppose that we can write the differential equation as

$$y' = F\left(\frac{y}{x}\right).$$

Here we try the substitutions

$$v = \frac{y}{x} \qquad \text{and therefore} \qquad y' = v + xv'.$$

We note that the equation is transformed into

$$v + xv' = F(v) \qquad \text{or} \qquad xv' = F(v) - v \qquad \text{or} \qquad \frac{v'}{F(v) - v} = \frac{1}{x}.$$

Hence an implicit solution is

$$\int \frac{1}{F(v) - v}\, dv = \ln|x| + C.$$

Example 39. *Solve*

$$x^2 y' = y^2 + xy, \qquad y(1) = 1.$$

Solution: We put the equation into the form $y' = ({}^{y}\!/\!{}_{x})^2 + {}^{y}\!/\!{}_{x}$. We substitute $v = {}^{y}\!/\!{}_{x}$ to get the separable equation

$$xv' = v^2 + v - v = v^2,$$

which has a solution

$$\begin{aligned} \int \frac{1}{v^2}\, dv &= \ln|x| + C, \\ \frac{-1}{v} &= \ln|x| + C, \\ v &= \frac{-1}{\ln|x| + C}. \end{aligned}$$

We unsubstitute

$$\begin{aligned} \frac{y}{x} &= \frac{-1}{\ln|x| + C}, \\ y &= \frac{-x}{\ln|x| + C}. \end{aligned}$$

We want $y(1) = 1$, so

$$1 = y(1) = \frac{-1}{\ln|1| + C} = \frac{-1}{C}.$$

Thus $C = -1$ and the solution we are looking for is

$$y = \frac{-x}{\ln|x| - 1}.$$

`firstOrder/practice/substitution-practice1.tex`

Practice for Substitution

Why?

Hint: Answers need not always be in closed form.

Exercise 205 *Solve* $y' + y(x^2 - 1) + xy^6 = 0$, *with* $y(1) = 1$.

Exercise 206 *Solve* $xy' + y + y^2 = 0$, $y(1) = 2$.

Exercise 207 *Solve* $2yy' + 1 = y^2 + x$, *with* $y(0) = 1$.

Exercise 208 *Solve* $xy' + y + x = 0$, $y(1) = 1$.

Exercise 209 *Solve* $y' + xy = y^4$, *with* $y(0) = 1$.

Exercise 210 *Solve* $y' + 3y = 2xy^4$.

Exercise 211 *Solve* $xy' - 2y = (3x^2 - x^{-3})y^5$ *with* $y(1) = 2$.

Exercise 212 *Solve* $y' + 5y = \dfrac{e^{2x}}{y^2}$.

Exercise 213 *Solve* $y^2y' = y^3 - 3x$, $y(0) = 2$.

Exercise 214 *Solve* $yy' + x = \sqrt{x^2 + y^2}$. *(Hint: What is* $\dfrac{d}{dx}(x^2 + y^2)$*)*

Exercise 215 *Solve* $y' = (x + y - 1)^2$.

Exercise 216 *Solve* $y' = \dfrac{x^2 - y^2}{xy}$, *with* $y(1) = 2$.

Author(s): Matthew Charnley and Jason Nowell

Exercise 217 *Solve* $2yy' = e^{y^2-x^2} + 2x$.

Exercise 218 *Consider the DE*

$$\frac{dy}{dt} = \left(y - \frac{1}{t}\right)^2 - \frac{1}{t^2}. \tag{15}$$

a) *Explain why* (15) *is not a linear equation.*

b) *Use a Bernoulli substitution to solve* (15).

First order linear PDE

We discuss First order linear PDE

We begin this chapter with an introduction to PDE in general, befcre moving on to techniques involving the Fourier Series discussed in Chapter **??**.

Consider the equation

$$a(x,t)\,u_x + b(x,t)\,u_t + c(x,t)\,u = g(x,t), \qquad u(x,0) = f(x), \qquad -\infty < x < \infty, \quad t > 0,$$

where $u(x,t)$ is a function of x and t. The *initial condition* $u(x,0) = f(x)$ is now a function of x rather than just a number. In these problems, it is useful to think of x as position and t as time. The equation describes the evolution of a function of x as time goes on. Below, the coefficients a, b, c, and the function g are mostly going to be constant or zero. The method we describe works with nonconstant coefficients, although the computations may get difficult quickly.

This method we use is the *method of characteristics*. The idea is that we find lines along which the equation is an ODE that we solve. We will see this technique again for second order PDE when we encounter the wave equation in § **??**.

Example 40. *Consider the equation*

$$u_t + \alpha u_x = 0, \qquad u(x,0) = f(x).$$

This particular equation, $u_t + \alpha u_x = 0$, is called the transport equation*.*

Solution: The data will propagate along curves called characteristics. The idea is to change to the so-called *characteristic coordinates*. If we change to these coordinates, the equation simplifies. The change of variables for this equation is

$$\xi = x - \alpha t, \qquad s = t.$$

Let's see what the equation becomes. Remember the chain rule in several variables.

$$\begin{aligned} u_t &= u_\xi \xi_t + u_s s_t = -\alpha u_\xi + u_s, \\ u_x &= u_\xi \xi_x + u_s s_x = u_\xi. \end{aligned}$$

The equation in the coordinates ξ and s becomes

$$\underbrace{(-\alpha u_\xi + u_s)}_{u_t} + \alpha \underbrace{(u_\xi)}_{u_x} = 0,$$

or in other words

$$u_s = 0.$$

That is trivial to solve. Treating ξ as simply a parameter, we have obtained the ODE $\dfrac{du}{ds} = 0$.

The solution is a function that does not depend on s (but it does depend on ξ). That is, there is some function A such that

$$u = A(\xi) = A(x - \alpha t).$$

The initial condition says that:

$$f(x) = u(x,0) = A(x - \alpha 0) = A(x),$$

so $A = f$. In other words,

$$u(x,t) = f(x - \alpha t).$$

Everything is simply moving right at speed α as t increases. The curve given by the equation

$$\xi = \text{constant}$$

is called the characteristic. See **Normally a reference to a previous figure goes here.**. In this case, the solution does not change along the characteristic.

Learning outcomes: Identify first order linear PDE Use the method of characteristics to solve first order PDE.
Author(s): Matthew Charnley and Jason Nowell

In the (x,t) coordinates, the characteristic curves satisfy $t = \frac{1}{\alpha}(x-\xi)$, and are in fact lines. The slope of characteristic lines is $\frac{1}{\alpha}$, and for each different ξ we get a different characteristic line.

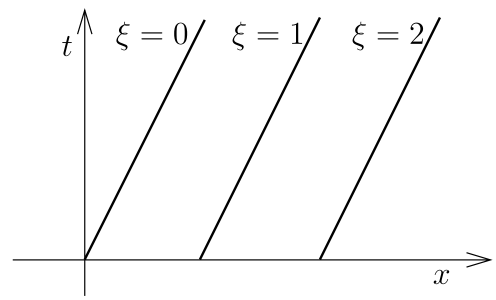


Figure 34: Characteristic curves.

We see why $u_t + \alpha u_x = 0$ is called the transport equation: everything travels at some constant speed. Sometimes this is called *convection*. An example application is material being moved by a river where the material does not diffuse and is simply carried along. In this setup, x is the position along the river, t is the time, and $u(x,t)$ the concentration the material at position x and time t. See **Normally a reference to a previous figure goes here.** for an example.

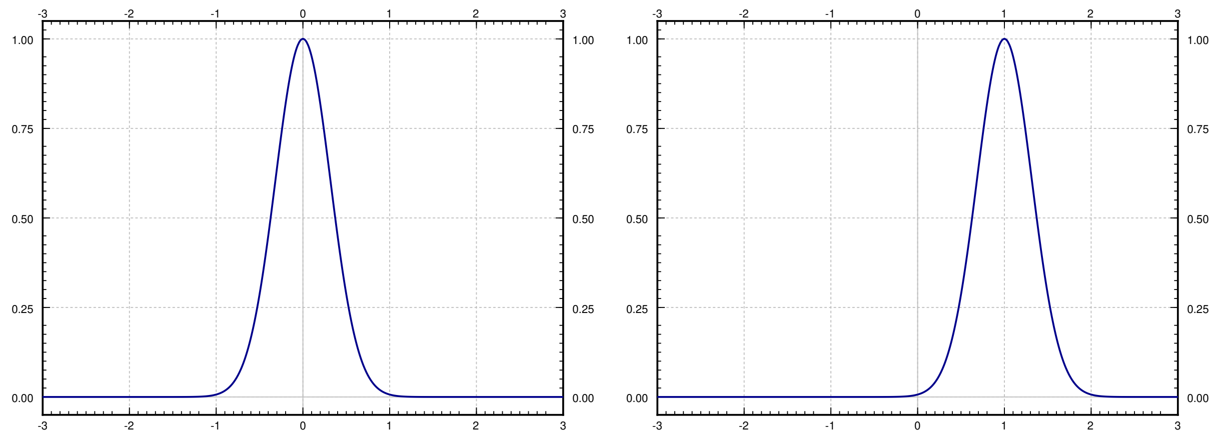


Figure 35: Example of "transport" in $u_t - u_x = 0$ (that is, $\alpha = 1$) where the initial condition $f(x)$ is a peak at the origin. On the left is a graph of the initial condition $u(x,0)$. On the right is a graph of the function $u(x,1)$, that is at time $t = 1$. Notice it is the same graph shifted one unit to the right.

We use similar idea in the more general case:

$$au_x + bu_t + cu = g, \qquad u(x,0) = f(x).$$

We change coordinates to the characteristic coordinates. Let us call these coordinates (ξ, s). These are coordinates where $au_x + bu_t$ becomes differentiation in the s variable.

Along the characteristic curves (where ξ is constant), we get a new ODE in the s variable. In the transport equation, we got the simple $\frac{du}{ds} = 0$. In general, we get the linear equation

$$\frac{du}{ds} + cu = g. \tag{16}$$

We think of everything as a function of ξ and s, although we are thinking of ξ as a parameter rather than an independent variable. So the equation is an ODE. It is a linear ODE that we can solve using the integrating factor.

To find the characteristics, think of a curve given parametrically $\big(x(s), t(s)\big)$. We try to have the curve satisfy

$$\frac{dx}{ds} = a, \qquad \frac{dt}{ds} = b.$$

Why? Because when we think of x and t as functions of s we find, using the chain rule,

$$\frac{du}{ds} + cu = \underbrace{\left(u_x \frac{dx}{ds} + u_t \frac{dt}{ds}\right)}_{\frac{du}{ds}} + cu = au_x + bu_t + cu = g.$$

So we get the ODE (16), which then describes the value of the solution u of the PDE along this characteristic curve. It is also convenient to make sure that $s = 0$ corresponds to $t = 0$, that is $t(0) = 0$. It will be convenient also for $x(0) = \xi$. See **Normally a reference to a previous figure goes here.**.

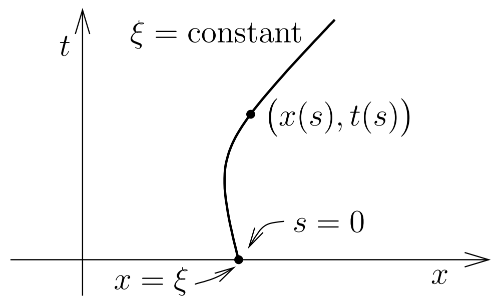


Figure 36: General characteristic curve.

Example 41. *Consider*

$$u_x + u_t + u = x, \qquad u(x,0) = e^{-x^2}.$$

Solution: We find the characteristics, that is, the curves given by

$$\frac{dx}{ds} = 1, \qquad \frac{dt}{ds} = 1.$$

So

$$x = s + c_1, \qquad t = s + c_2,$$

for some c_1 and c_2. At $s = 0$ we want $t = 0$, and x should be ξ. So we let $c_1 = \xi$ and $c_2 = 0$:

$$x = s + \xi, \qquad t = s.$$

The ODE is $\dfrac{du}{ds} + u = x$, and $x = s + \xi$. So, the ODE to solve along the characteristic is

$$\frac{du}{ds} + u = s + \xi.$$

The general solution of this equation, treating ξ as a parameter, is $u = Ce^{-s} + s + \xi - 1$, for some constant C. At $s = 0$, our initial condition is that u is $e^{-\xi^2}$, since at $s = 0$ we have $x = \xi$. Given this initial condition, we find $C = e^{-\xi^2} - \xi + 1$. So,

$$\begin{aligned} u &= \left(e^{-\xi^2} - \xi + 1\right)e^{-s} + s + \xi - 1 \\ &= e^{-\xi^2 - s} + (1-\xi)e^{-s} + s + \xi - 1. \end{aligned}$$

Substitute $\xi = x - t$ and $s = t$ to find u in terms of x and t:

$$\begin{aligned} u &= e^{-\xi^2 - s} + (1-\xi)e^{-s} + s + \xi - 1 \\ &= e^{-(x-t)^2 - t} + (1 - x + t)e^{-t} + x - 1. \end{aligned}$$

See **Normally a reference to a previous figure goes here.** for a plot of $u(x,t)$ as a function of two variables.

When the coefficients are not constants, the characteristic curves are not going to be straight lines anymore.

Example 42. *Consider the following variable coefficient equation:*

$$xu_x + u_t + 2u = 0, \qquad u(x,0) = \cos(x).$$

Solution: We find the characteristics, that is, the curves given by

$$\frac{dx}{ds} = x, \qquad \frac{dt}{ds} = 1.$$

So

$$x = c_1 e^s, \qquad t = s + c_2.$$

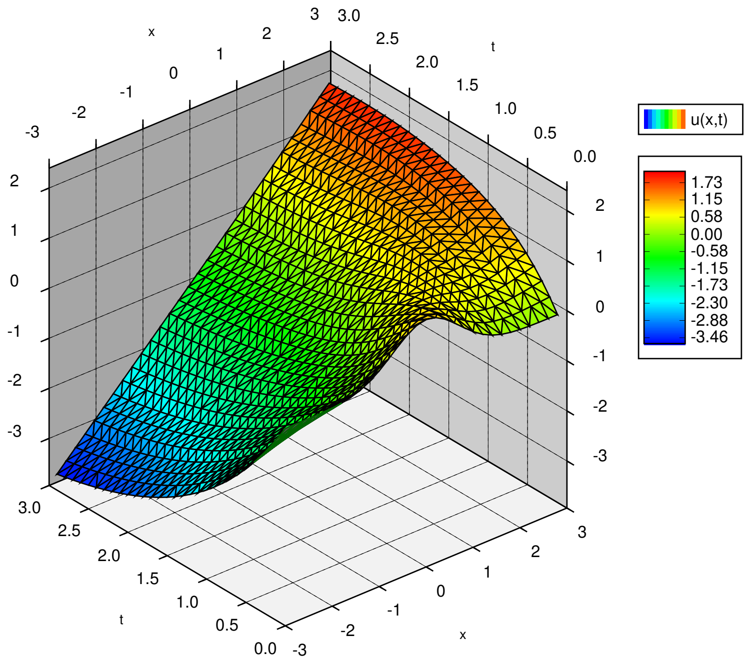


Figure 37: Plot of the solution $u(x,t)$ to $u_x + u_t + u = x$, $u(x,0) = e^{-x^2}$.

At $s = 0$, we wish to get the line $t = 0$, and x should be ξ. So

$$x = \xi e^s, \qquad t = s.$$

OK, the ODE we need to solve is

$$\frac{du}{ds} + 2u = 0.$$

This is for a fixed ξ. At $s = 0$, we should get that u is $\cos(\xi)$, so that is our initial condition. Consequently,

$$u = e^{-2s}\cos(\xi) = e^{-2t}\cos(xe^{-t}).$$

We make a few closing remarks. One thing to keep in mind is that we would get into trouble if the coefficient in front of u_t, that is the b, is ever zero. Let us consider a quick example of what can go wrong:

$$u_x + u = 0, \qquad u(x,0) = \sin(x).$$

This problem has no solution. If we had a solution, it would imply that $u_x(x,0) = \cos(x)$, but $u_x(x,0) + u(x,0) = \cos(x) + \sin(x) \neq 0$. The problem is that the characteristic curve is now the line $t = 0$, and the solution is already provided on that line!

As long as b is nonzero, it is convenient to ensure that b is positive by multiplying by -1 if necessary, so that positive s means positive t.

Another remark is that if a or b in the equation are variable, the computations can quickly get out of hand, as the expressions for the characteristic coordinates become messy and then solving the ODE becomes even messier. In the examples above, b was always 1, meaning we got $s = t$ in the characteristic coordinates. If b is not constant, your expression for s will be more complicated.

Finding the characteristic coordinates is really a system of ODE in general if a depends on t or if b depends on x. In that case, we would need techniques of systems of ODE to solve, see chapter **??** or chapter **??**. In general, if a and b are not linear functions or constants, finding closed form expressions for the characteristic coordinates may be impossible.

Finally, the method of characteristics applies to nonlinear first order PDE as well. In the nonlinear case, the characteristics depend not only on the differential equation, but also on the initial data. This leads to not only more difficult computations, but also the formation of singularities where the solution breaks down at a certain point in time. An example application where first order nonlinear PDE come up is traffic flow theory, and you have probably experienced the formation of singularities: traffic jams. But we digress.

firstOrder/practice/firstPDE-practice1.tex

Practice for first PDE

Why?

Exercise 219 *Solve*

a) $u_t + 9u_x = 0, \quad u(x,0) = \sin(x),$

b) $u_t - 8u_x = 0, \quad u(x,0) = \sin(x),$

c) $u_t + \pi u_x = 0, \quad u(x,0) = \sin(x),$

d) $u_t + \pi u_x + u = 0, \quad u(x,0) = \sin(x).$

Exercise 220 *Solve*

a) $u_t - 5u_x = 0,\ u(x,0) = \dfrac{1}{1+x^2},$

b) $u_t + 2u_x = 0,\ u(x,0) = \cos(x).$

Exercise 221 *Solve* $u_t + 3u_x = 1,\ u(x,0) = x^2.$

Exercise 222 *Solve* $u_x + u_t + tu = 0,\ u(x,0) = \cos(x).$

Exercise 223 *Solve* $u_t + 3u_x = x,\ u(x,0) = e^x.$

Exercise 224 *Solve* $u_x + u_t + xu = 0,\ u(x,0) = \cos(x).$

Exercise 225 *Solve* $u_x + u_t = 5,\ u(x,0) = x.$

Exercise 226

a) *Find the characteristic coordinates for the following equations:*
   1) $u_x + u_t + u = 1,\ \ u(x,0) = \cos(x),$ *2)* $2u_x + 2u_t + 2u = 2,\ \ u(x,0) = \cos(x).$

b) *Solve the two equations using the coordinates.*

c) *Explain why you got the same solution, although the characteristic coordinates you found were different.*

Exercise 227 *Solve* $(1+x^2)u_t + x^2u_x + e^x u = 0,\ u(x,0) = 0.$ *Hint: Think a little out of the box.*

Author(s): Matthew Charnley and Jason Nowell

Chapter Review

Review for Chapter on First Order ODEs.

Exercise 228 *Bernoulli the snail slithers along a tabletop measuring 20 inches across, starting at one edge with a remarkable speed of 4 inches per hour. Bernoulli is afraid of falling off the other edge of the table—as you would be too if you were a snail. Accordingly, he slows down so that the speed of the snail is proportional to the square of the distance remaining. How long will it take Bernoulli to get halfway across the table?*

Author(s): Matthew Charnley and Jason Nowell

Second order linear ODEs

We discuss Second order linear ODEs

The general second order ordinary differential equation is of the form

$$y'' = F(x, y, y')$$

for F an arbitrary function of three variables. As with first order equations, if the function F is not in a nice or simple form, there really isn't a hope to find a solution for this. For second order equations, we need to be even more specific about the structure of these equations in order to find solutions than we did for first order.

Definition 10. *The general* second order linear differential equation *is of the form*

$$A(x)y'' + B(x)y' + C(x)y = F(x).$$

This equation can be written in standard form *by dividing through by $A(x)$ to get*

$$y'' + p(x)y' + q(x)y = f(x), \tag{17}$$

where $p(x) = {}^{B(x)}/_{A(x)}$, $q(x) = {}^{C(x)}/_{A(x)}$, and $f(x) = {}^{F(x)}/_{A(x)}$.

The word *linear* means that the equation contains no powers nor functions of y, y', and y''. In the special case when $f(x) = 0$, we have a so-called *homogeneous* equation

$$y'' + p(x)y' + q(x)y = 0. \tag{18}$$

We have already seen some second order linear homogeneous equations.

$$\begin{aligned} &y'' + k^2 y = 0 && \text{Two solutions are:} \quad y_1 = \cos(kx), \quad y_2 = \sin(kx). \\ &y'' - k^2 y = 0 && \text{Two solutions are:} \quad y_1 = e^{kx}, \quad y_2 = e^{-kx}. \end{aligned}$$

With the examples above, we were able to find solutions. However, notice that these equations don't have functions of x as coefficients of the y term. This means they are constant coefficient equations. It turns out that one of the few ways we can have a guaranteed method for finding solutions to these equation is if they have constant coefficients. For first order, we had a method for every linear equation, but for second order, we only have a formulaic method for constant coefficient and homogeneous linear equations.

If we know two solutions of a linear homogeneous equation, we know many more of them.

Theorem 4. *Superposition Suppose y_1 and y_2 are two solutions of the homogeneous equation* (18)*. Then*

$$y(x) = C_1 y_1(x) + C_2 y_2(x),$$

also solves (18) *for arbitrary constants C_1 and C_2.*

That is, we can add solutions together and multiply them by constants to obtain new and different solutions. We call the expression $C_1y_1 + C_2y_2$ a *linear combination* of y_1 and y_2. Let us prove this theorem; the proof is very enlightening and illustrates how linear equations work.

Proof: Let $y = C_1y_1 + C_2y_2$. Then

$$\begin{aligned} y'' + py' + qy &= (C_1y_1 + C_2y_2)'' + p(C_1y_1 + C_2y_2)' + q(C_1y_1 + C_2y_2) \\ &= C_1y_1'' + C_2y_2'' + C_1py_1' + C_2py_2' + C_1qy_1 + C_2qy_2 \\ &= C_1(y_1'' + py_1' + qy_1) + C_2(y_2'' + py_2' + qy_2) \\ &= C_1 \cdot 0 + C_2 \cdot 0 = 0. \quad \blacksquare \end{aligned}$$

Learning outcomes: Identify the general second order linear differential equation Determine the characteristic equation for constant coefficient equations Find the general solution for constant coefficient equations in the real and distinct roots case Determine if two functions are linearly independent.

Author(s): Matthew Charnley and Jason Nowell

The proof becomes even simpler to state if we use the operator notation. An *operator* is an object that eats functions and spits out functions (kind of like what a function is, but a function eats numbers and spits out numbers). Define the operator L by

$$L[y] = y'' + py' + qy.$$

The differential equation now becomes $L[y] = 0$. The operator (and the equation) L being *linear* means that $L[C_1 y_1 + C_2 y_2] = C_1 L[y_1] + C_2 L[y_2]$. The proof above becomes

$$L[y] = L[C_1 y_1 + C_2 y_2] = C_1 L[y_1] + C_2 L[y_2] = C_1 \cdot 0 + C_2 \cdot 0 = 0.$$

Exercise 229 *This fact does not hold if the equation is non-linear. Show that $y_1(t) = e^t$ and $y_2(t) = 1$ solve*

$$y'' = \sqrt{y \cdot y'}$$

but $y(t) = e^t + 1$ does not.

Two different solutions to the second equation $y'' - k^2 y = 0$ are $y_1 = \cosh(kx)$ and $y_2 = \sinh(kx)$. Let us remind ourselves of the definition, $\cosh x = \dfrac{e^x + e^{-x}}{2}$ and $\sinh x = \dfrac{e^x - e^{-x}}{2}$. Therefore, these are solutions by superposition as they are linear combinations of the two exponential solutions.

The functions sinh and cosh are sometimes more convenient to use than the exponential. Let us review some of their properties:

$$\cosh 0 = 1, \qquad \sinh 0 = 0,$$

$$\frac{d}{dx}\Big[\cosh x\Big] = \sinh x, \qquad \frac{d}{dx}\Big[\sinh x\Big] = \cosh x,$$

$$\cosh^2 x - \sinh^2 x = 1.$$

Exercise 230 *Derive these properties using the definitions of* sinh *and* cosh *in terms of exponentials.*

Intial Value Problems

For first order equations, a lot of problems were stated as Initial Value Problems, containing both a differential equation and an initial condition of the value of y at some point x_0. What do these initial condition(s) look like for second order equations?

Example 43. *Solve the second-order differential equation*

$$y'' = x.$$

Solution: We can attempt to find a solution to this problem by integrating both sides twice. A first integration gives

$$y' = \frac{x^2}{2} + C$$

and a second integration leads to

$$y = \frac{x^3}{6} + Cx + D$$

for any two constants C and D. We can check that differentiating this y function twice gives us back the function x that we wanted.

In the previous example, we ended up with two unknown constants in our answer, whereas for first order equations, we only had one. In order to specify these two constants, we will need to give two additional facts about this function. This could be the value of the function at two points, but more traditionally, it is given as the value of the function y and its first derivative y' at a value x_0. Fairly often, this value x_0 is 0, but it could be any other number.

Example 44. *Solve the initial value problem*

$$y'' = x, \qquad y(1) = 2,\ y'(1) = 3$$

Solution: We previously found our solution with unknown constants as

$$y = \frac{x^3}{6} + Cx + D$$

and also found that

$$y' = \frac{x^2}{2} + C.$$

To find the values of C and D, we need to plug in the two initial conditions into their corresponding functions. The initial value of the derivative gives that

$$3 = y'(1) = \frac{1^2}{2} + C = C + \frac{1}{2}$$

so that we have $C = \frac{5}{2}$. We can then use the initial value of y, along with this C value, to conclude that

$$2 = y(1) = \frac{1^3}{6} + \frac{5}{2}(1) + D = \frac{1}{6} + \frac{5}{2} + D = \frac{16}{6} + D.$$

Solving this out gives that $D = -\frac{4}{6} = -\frac{2}{3}$. Putting these constants in gives that the solution to the initial value problem is

$$y = \frac{x^3}{6} + \frac{5}{2}x - \frac{2}{3}.$$

For first-order equations, we have theorems that told us that solutions existed and were unique, at least on small intervals. Linear first-order equations in particular had a very nice existence and uniqueness theorem (Theorem **??**), guaranteeing existence on a full interval wherever the coefficient functions are continuous. Linear second-order equations have an existence and uniqueness theorem that gives the same type of result when the initial condition is stated properly.

Theorem 5. *Existence and uniqueness Suppose p, q, f are continuous functions on some interval I, a is a number in I, and a, b_0, b_1 are constants. The equation*

$$y'' + p(x)y' + q(x)y = f(x),$$

has exactly one solution $y(x)$ defined on the same interval I satisfying the initial conditions

$$y(a) = b_0, \qquad y'(a) = b_1.$$

For example, the equation $y'' + k^2y = 0$ with $y(0) = b_0$ and $y'(0) = b_1$ has the solution

$$y(x) = b_0 \cos(kx) + \frac{b_1}{k} \sin(kx).$$

The equation $y'' - k^2y = 0$ with $y(0) = b_0$ and $y'(0) = b_1$ has the solution

$$y(x) = b_0 \cosh(kx) + \frac{b_1}{k} \sinh(kx).$$

Using cosh and sinh in this solution allows us to solve for the initial conditions in a cleaner way than if we have used the exponentials.

As it did for first order equations, this theorem tells us what the proper form is for initial value problems for second order equations. The take-away here is that in order to fully specify a solution to an initial value problem, a second order equation requires two initial conditions. They are usually given in the form $y(a)$ and $y'(a)$, but could be given as $y(a_1)$ and $y(a_2)$ in other applications. In any case, two pieces of information are needed to determine a problem of second order, where we only needed one for first order.

Constant Coefficient Equations - Real and Distinct Roots

Now we want to try to solve some of these equations. As discussed earlier in this section, there is no explicit solution method possible for second order equations. However, if we restrict to a very simple case (which is also one that shows up frequently in physical systems) we can start to develop a method for solving these equations. The type of equation we restrict to is linear and constant coefficient equations. *Constant coefficients* means that the functions in front of y'', y', and y are constants, they do not depend on x. The most general second order, linear, constant coefficient equation is

$$ay'' + by' + cy = g(x)$$

for real constants a, b, c and an arbitrary function $g(x)$. We will study the solution of nonhomogeneous equations (with $g(x) \neq 0$) in § . We will first focus on finding general solutions to homogeneous equations, which are of the form

$$ay'' + by' + cy = 0.$$

Consider the problem

$$y'' - 6y' + 8y = 0.$$

This is a second order linear homogeneous equation with constant coefficients, so it fits the type of equation where we want to hunt for solutions. To guess a solution, think of a function that stays essentially the same when we differentiate it, so that we can take the function and its derivatives, add some multiples of these together, and end up with zero. Yes, we are talking about the exponential.

Let us try[1] a solution of the form $y = e^{rx}$. Then $y' = re^{rx}$ and $y'' = r^2e^{rx}$. Plug in to get

$$\begin{aligned} y'' - 6y' + 8y &= 0, \\ \underbrace{r^2e^{rx}}_{y''} - 6\underbrace{re^{rx}}_{y'} + 8\underbrace{e^{rx}}_{y} &= 0, \\ r^2 - 6r + 8 &= 0 \qquad \text{(divide through by } e^{rx}\text{)}, \\ (r-2)(r-4) &= 0. \end{aligned}$$

Hence, if $r = 2$ or $r = 4$, then e^{rx} is a solution. So let $y_1 = e^{2x}$ and $y_2 = e^{4x}$.

Exercise 231 *Check that y_1 and y_2 are solutions.*

So we have found two solutions to this differential equation! That's great, but there may be a few concerning ideas at this point:

(a) Did we just get lucky with this particular equation?

(b) How do we know that there aren't other solutions that aren't of the form e^{rx}? We made that assumption, so we could have missed something.

The second point comes back to the existence and uniqueness theorem. This differential equation satisfies the conditions of the existence and uniqueness theorem. That means that as long as we find *a* solution that can meet any initial condition, then we know that the solution we have found is the *only* solution. We have not yet verified the part about meeting initial conditions yet (that's coming later), but once we do, we'll know that making this assumption is completely fine, because it got us to a solution that works, and the uniqueness theorem tells us that this is the only solution.

For the first point, let's try to generalize the calculation we did above into a method that will work for more equations. Suppose that we have an equation

$$ay'' + by' + cy = 0, \tag{19}$$

where a, b, c are constants. We can take our same assumption that the solution is of the form $y = e^{rx}$ to obtain

$$ar^2e^{rx} + bre^{rx} + ce^{rx} = 0.$$

[1] Making an educated guess with some parameters to solve for is such a central technique in differential equations, that people sometimes use a fancy name for such a guess: *ansatz*, German for "initial placement of a tool at a work piece". Yes, the Germans have a word for that.

Divide by e^{rx} to obtain the so-called *characteristic equation* of the ODE:

$$ar^2 + br + c = 0.$$

Solve for the r by using the quadratic formula.

$$r_1, r_2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

There are three cases that can arise based on this equation.

(a) If $b^2 - 4ac > 0$, then we have r_1 and r_2 as two real roots to the equation. This is the same as the example above, and we get $e^{r_1 x}$ and $e^{r_2 x}$ as two solutions. This is the larger class of problems to which this exact process applies.

(b) If $b^2 - 4ac < 0$, then r_1 and r_2 are complex numbers. We can still use $e^{r_1 x}$ and $e^{r_2 x}$ as solutions, but this runs into some issues, which will be addressed in Section .

(c) If $b^2 - 4ac = 0$, then we only get one root, since $r_1 = r_2$. We do get that $e^{r_1 x}$ as a solution, but that's all we get. This is another issue, which is addressed in Section .

So, as long as we have $b^2 - 4ac > 0$, this method will work to give us two solutions to this differential equation.

Example 45. *Find two values of r so that e^{rx} is a solution to*

$$y'' + 3y' - 10y = 0$$

Our first step is to find the characteristic equation by plugging e^{rx} into the equation. This gives that

$$r^2 + 3r - 10 = 0$$

This polynomial factors as $(r-2)(r+5)$, so we know that values of $r = 2$ and $r = -5$ will work. This means (check this!) that e^{2x} and e^{-5x} solve this differential equation.

Linear Independence

Since e^{2x} and e^{-5x} solve the linear differential equation in the previous example, we know that superposition applies, so that $C_1 e^{2x} + C_2 e^{-5x}$ solves the differential equation for any C_1 and C_2. The last thing to check is that we can pick C_1 and C_2 in order to meet any initial condition that we want. If this is possible, then we know that our method using the characteristic equation to find e^{2x} and e^{-5x} as solutions was enough to always solve this problem. The end of this argument is done using the existence and uniqueness theorem as described previously.

Let's work this out. Assume that we are given b_0 and b_1 and want to solve the initial value problem

$$y'' + 3y' - 10y = 0 \qquad y(0) = b_0,\ y'(0) = b_1.$$

We want to do this by picking C_1 and C_2 in the expression $y = C_1 e^{2x} + C_2 e^{-5x}$ Since

$$y' = 2C_1 e^{2x} - 5C_2 e^{-5x}$$

we can plug zero into this equation and the equation for y to get that we would need to have

$$\begin{aligned} b_0 = y(0) = C_1 + C_2 \\ b_1 = y'(0) = 2C_1 - 5C_2 \end{aligned}.$$

We can solve this system of equations by elimination. Multiplying the first equation by 5 adding them together gives

$$5b_0 + b_1 = 7C_1$$

so that

$$C_1 = \frac{5b_0 + b_1}{7}.$$

We can then compute the value of C_2 as

$$C_2 = b_0 - C_1 = b_0 - \frac{5b_0 + b_1}{7} = \frac{2b_0 - b_1}{7}.$$

Therefore, we can appropriate values of C_1 and C_2 that will meet the initial conditions for arbitrary values b_0 and b_1. This is great! This means that our method of finding solutions was sufficient for this problem.

Let's look at this situation in more generality. Assume that we have two solutions y_1 and y_2 that solve a second order linear, homogeneous differential equation, and we want to know if $C_1y_1 + C_2y_2$ can meet any initial condition for this problem. We have two unknowns and two equations ($y(x_0)$ and $y'(x_0)$ for some value x_0), so it should work out.

We can carry out the same steps as above. If we have initial conditions $y(x_0) = b_0$ and $y'(x_0) = b_1$, we want to satisfy

$$\begin{aligned} b_0 &= y(x_0) = C_1y_1(x_0) + C_2y_2(x_0) \\ b_1 &= y'(x_0) = C_1y_1'(x_0) + C_2y_2'(x_0)' \end{aligned}$$

which we get by taking the derivative of $y(x) = C_1y_1(x) + C_2y_2(x)$ and plugging in x_0. We will again use elimination to solve this. We can multiply the first equation by $y_1'(x_0)$, multiply the second by $y_1(x_0)$, and subtract them. This will cancel out the C_1 term, leaving us with

$$b_0y_1'(x_0) - b_1y_1(x_0) = C_2(y_1'(x_0)y_2(x_0) - y_1(x_0)y_2'(x_0)).$$

We want to solve for C_2 here, and once we do that, solving for C_1 happens by plugging back into one of the original equations. Most of the time, this will be completely fine, but there's one issue left. We can't divide by zero. So to be able to solve these equations for C_1 and C_2, we need to know that

$$y_1'(x_0)y_2(x_0) - y_1(x_0)y_2'(x_0) \neq 0. \tag{20}$$

The left side of this equation is often called the *Wronskian* of the functions y_1 and y_2 at the point x_0. In general, the Wronskian is the function $y_1'(x)y_2(x) - y_2'(x)y_1(x)$ for two solutions to a second order differential equation. This relation (the Wronskian being non-zero) tells us that the two solutions y_1 and y_2 are different enough to allow us to meet every initial condition for the differential equation. This condition is so important to the study of second order linear equations that we give it a name. We say that two solutions y_1 and y_2 are *linearly independent at* x_0 if (20) holds, that is, if the Wronskian of the solutions is non-zero at that point. For two solutions of a differential equation (which is more specific than just having two random functions), two solutions being linearly independent is equivalent to 20 holding for any[2] value x_0 where they are defined. Our work and calculations above leads to the following theorem:

Theorem 6. *Let p, q be continuous functions. Let y_1 and y_2 be two linearly independent solutions to the homogeneous equation* (18). *Then every other solution is of the form*

$$y = C_1y_1 + C_2y_2$$

for some constants C_1 and C_2. That is, $y = C_1y_1 + C_2y_2$ is the general solution.

Note that this theorem works for all linear homogeneous equations, not just constant coefficients ones. However, the methods that we have described here (and will in future sections) for *finding* these solutions will generally only work for constant coefficient equations.

This idea of linear independence can also be expressed in a different way: two solutions y_1 and y_2 are linearly independent if only way to make the expression

$$c_1y_1 + c_2y_2 = 0$$

is by setting both $c_1 = 0$ and $c_2 = 0$. This comes from the idea of linear independence from linear algebra (see Chapter **??**) and uniqueness of solutions to differential equations. If there are such constants, we can also rearrange the equation to give

$$y_1 = -\frac{c_2}{c_1}y_2$$

which says that y_1 is a constant multiple of y_2, which holds for all values of x. Thus, if we have y_1 and y_2, and there is no constant A so that $y_1 = Ay_2$, then these functions are linearly independent.

Example 46. *Find the general solution of the differential equation $y'' + y = 0$.*

Solution: One of the four fundamental equations in § showed that the two functions $y_1 = \sin x$ and $y_2 = \cos x$ are solutions to the equation $y'' + y = 0$. It is not hard to see that sine and cosine are not constant multiples of each other. If $\sin x = A\cos x$ for some **constant** A, we let $x = 0$ and this would imply $A = 0$. But then $\sin x = 0$ for all x, which is

[2] Abel's Theorem, another theoretical result, says that the Wronskian $y_1'y_2 - y_1y_2'$ is either always zero or never zero. That means that any one value can be checked to determine if two solutions are linearly independent. Picking 0 is usually a convenient choice.

preposterous. So y_1 and y_2 are linearly independent. We could also have checked this by taking derivatives and plugging in zero. Since

$$y_1(0) = 0 \quad y_1'(0) = 1 \quad y_2(0) = 1 \quad y_2'(0) = 0$$

we have that

$$y_1'(0)y_2(0) - y_1(0)y_2'(0) = (1)(1) - (0)(0) = 1 \neq 0$$

so these solutions are linearly independent. Hence,

$$y = C_1 \cos x + C_2 \sin x$$

is the general solution to $y'' + y = 0$.

For two functions, checking linear independence is rather simple. Let us see another example using non-constant coefficient equations. Consider $y'' - 2x^{-2}y = 0$. Then $y_1 = x^2$ and $y_2 = 1/x$ are solutions. To see that they are linearly indepedent, suppose one is a multple of the other: $y_1 = Ay_2$, we just have to find out that A cannot be a constant. In this case we have $A = y_1/y_2 = x^3$, this most decidedly not a constant. So $y = C_1x^2 + C_2 1/x$ is the general solution.

Now, back to our discussion of constant coefficient equations. If $b^2 - 4ac > 0$, then we have two distinct real roots r_1 and r_2, giving rise to solutions of the form $y_1(x) = e^{r_1x}$ and $y_2(x) = e^{r_2x}$. Using condition 20 with $x_0 = 0$, we compute

$$y_1'(0)y_2(0) - y_1(0)y_2'(0) = (r_1)(1) - (1)(r_2) = r_1 - r_2.$$

Since $r_1 \neq r_2$, this expression is not zero, so the two solutions are linearly independent. Therefore, in this case, we know that the general solution will be

$$y = C_1e^{r_1x} + C_2e^{r_2x}.$$

Using the other formulation of linear independence of two functions, we would need to show that there is no constant A so that

$$e^{r_1x} = Ae^{r_2x}.$$

Since this can be rewritten as $A = e^{(r_1-r_2)x}$ and we know that $r_1 \neq r_2$, this is not a constant, so we again know that these functions are linearly independent and give rise to a general solution.

Example 47. *Solve the initial value problem*

$$y'' + 2y' - 3y = 0 \qquad y(0) = 2,\ y'(0) = 1.$$

Solution: To start, we find the characteristic equation of this differential equation and look for the roots. The characteristic equation here is

$$r^2 + 2r - 3 = 0$$

and this factors as $(r + 3)(r - 1) = 0$. Thus, the two roots are $r = 1$ and $r = -3$, so that the general solution (and we know it is the general solution because these are different exponents and so the solutions are linearly independent) is

$$y(x) = C_1e^x + C_2e^{-3x}.$$

In order to find the values of C_1 and C_2, we need to use the initial conditions. Plugging zero into $y(x)$ gives

$$y(0) = 2 = C_1 + C_2$$

and since the derivative $y'(x) = C_1e^x - 3C_2e^{-3x}$, the second condition gives that

$$y'(0) = 1 = C_1 - 3C_2.$$

Subtracting the second equation from the first gives that

$$1 = 4C_2$$

so that $C_2 = 1/4$ and $C_1 = 7/4$. Thus, the solution to the initial value problem is

$$y(x) = \frac{7}{4}e^x + \frac{1}{4}e^{-3x}.$$

In this second example, we solve a problem in the same way, but the roots of the characteristic equation do not work out as nicely. Even with that, the structure and process for the problem is identical to the previous example.

Example 48. *Solve the initial value problem*

$$y'' - 2y' - y = 0 \qquad y(0) = 2,\ y'(0) = 3.$$

Solution: We start by looking for the characteristic equation of this differential equation and finding its roots. The characteristic equation is

$$r^2 - 2r - 1 = 0$$

which has roots

$$r = \frac{2 \pm \sqrt{(-2)^2 - 4(1)(-1)}}{2} = \frac{2 \pm \sqrt{8}}{2} = 1 \pm \sqrt{2}.$$

There are two real and distinct roots, so we know that the two solutions $y_1(x) = e^{(1+\sqrt{2})x}$ and $y_2(x) = e^{(1-\sqrt{2})x}$ are linearly independent, so we have that the general solution to this problem is

$$y(x) = C_1 e^{(1+\sqrt{2})x} + C_2 e^{(1-\sqrt{2})x}.$$

Next, we need to find the constants C_1 and C_2 to meet the initial conditions. We can see that, by computing the first derivative,

$$\begin{aligned} y(x) &= C_1 e^{(1+\sqrt{2})x} + C_2 e^{(1-\sqrt{2})x}, \\ y'(x) &= (1+\sqrt{2})C_1 e^{(1+\sqrt{2})x} + (1-\sqrt{2})C_2 e^{(1-\sqrt{2})x}, \end{aligned}$$

and plugging in $x = 0$ gives that we want C_1 and C_2 to solve

$$\begin{aligned} 2 &= C_1 + C_2, \\ 3 &= (1+\sqrt{2})C_1 + (1-\sqrt{2})C_2. \end{aligned}$$

We can solve this by any method. One trick at the start is to subtract equation 1 from equation 2, giving that

$$\begin{aligned} 2 &= C_1 + C_2, \\ 1 &= \sqrt{2}C_1 - \sqrt{2}C_2, \end{aligned}$$

which can be rewritten as

$$\begin{aligned} 2 &= C_1 + C_2, \\ \frac{1}{\sqrt{2}} &= C_1 - C_2. \end{aligned}$$

Adding these equations together and dividing by 2 gives that

$$2C_1 = 2 + \frac{1}{\sqrt{2}}$$

so that $C_1 = 1 + \frac{1}{2\sqrt{2}}$, and since $C_1 + C_2 = 2$, we have that $C_2 = 1 - \frac{1}{2\sqrt{2}}$. Therefore, the solution to the desired initial value problem is

$$y(x) = \left(1 + \frac{1}{2\sqrt{2}}\right) e^{(1+\sqrt{2})x} + \left(1 - \frac{1}{2\sqrt{2}}\right) e^{(1-\sqrt{2})x}.$$

$\lrcorner$

`higherOrder/practice/secondOrderLinearODE-practice1.tex`

Practice for Second Order Linear ODEs

Why?

Exercise 232 *Show that* $y = e^x$ *and* $y = e^{2x}$ *are linearly independent.*

Exercise 233 *Are* $\sin(x)$ *and* e^x *linearly independent? Justify.*

Exercise 234 *Are* e^x *and* e^{x+2} *linearly independent? Justify.*

Exercise 235 *Guess a solution to* $y'' + y' + y = 5$.

Exercise 236 *Take* $y'' + 5y = 10x + 5$. *Find (guess!) a solution.*

Exercise 237 *Verify that* $y_1(t) = e^t \cos(2t)$ *and* $y_2(t) = e^t \sin(2t)$ *both solve* $y'' - 2y' + 5y = 0$. *Are these two solutions linearly independent? What does that mean about the general solution to* $y'' - 2y' + 5y = 0$*?*

Exercise 238 *Prove the superposition principle for nonhomogeneous equations. Suppose that* y_1 *is a solution to* $Ly_1 = f(x)$ *and* y_2 *is a solution to* $Ly_2 = g(x)$ *(same linear operator* L*). Show that* $y = y_1 + y_2$ *solves* $Ly = f(x) + g(x)$.

Exercise 239 *Determine the maximal interval of existence of the solution to the differential equation*

$$(t-5)y'' + \frac{1}{t+1}y' + e^t y = \frac{\cos(t)}{t^2+1}$$

with initial condition $y(3) = 8$. *What about if the initial condition is* $y(-3) = 4$*?*

Exercise 240 *For the equation* $x^2y'' - xy' = 0$, *find two solutions, show that they are linearly independent and find the general solution. Hint: Try* $y = x^r$.

Exercise 241 *Find the general solution to* $xy'' + y' = 0$. *Hint: It is a first order ODE in* y'.

Exercise 242 *Find the general solution of* $2y'' + 2y' - 4y = 0$.

Author(s): Matthew Charnley and Jason Nowell

Exercise 243 *Solve $y'' + 9y' = 0$ with $y(0) = 1$, $y'(0) = 1$.*

Exercise 244 *Find the general solution of $y'' + 9y' - 10y = 0$.*

Exercise 245 *Find the general solution to $y'' - 3y' - 4y = 0$.*

Exercise 246 *Find the general solution to $y'' + 6y' + 8y = 0$.*

Exercise 247 *Find the solution to $y'' - 3y' + 2y = 0$ with $y(0) = 3$ and $y'(0) = -1$.*

Exercise 248 *Find the solution to $y'' + y' - 12y = 0$ with $y(0) = 1$ and $y'(0) = -2$.*

Exercise 249 *Find the general solution to $y'' + 4y' + 2y = 0$.*

Exercise 250 *Find the solution to $2y'' + y' - 3y = 0$, $y(0) = a$, $y'(0) = b$.*

Exercise 251 *Find the solution to $y'' - (\alpha + \beta)y' + \alpha\beta y = 0$, $y(0) = a$, $y'(0) = b$, where α, β, a, and b are real numbers, and $\alpha \neq \beta$.*

Exercise 252 *Write down an equation (guess) for which we have the solutions e^x and e^{2x}. Hint: Try an equation of the form $y'' + Ay' + By = 0$ for constants A and B, plug in both e^x and e^{2x} and solve for A and B.*

Exercise 253 *Construct an equation such that $y = C_1 e^{3x} + C_2 e^{-2x}$ is the general solution.*

Exercise 254 *Give an example of a 2nd-order DE whose general solution is $y = c_1 e^{-2t} + c_2 e^{-4t}$.*

Equations of the form $ax^2y'' + bxy' + cy = 0$ are called *Euler's equations* or *Cauchy–Euler equations.* They are solved by trying $y = x^r$ and solving for r (assume that $x \geq 0$ for simplicity).

Exercise 255 *Suppose that $(b - a)^2 - 4ac > 0$.*

a) *Find a formula for the general solution of $ax^2y'' + bxy' + cy = 0$. Hint: Try $y = x^r$ and find a formula for r.*
b) *What happens when $(b - a)^2 - 4ac = 0$ or $(b - a)^2 - 4ac < 0$?*

We will revisit the case when $(b - a)^2 - 4ac < 0$ later.

Exercise 256 *Same equation as in* Exercise . *Suppose $(b - a)^2 - 4ac = 0$. Find a formula for the general solution of $ax^2y'' + bxy' + cy = 0$. Hint: Try $y = x^r \ln x$ for the second solution.*

Complex Roots and Euler's Formula

We discuss Complex Roots and Euler's Formula

The next case to consider for constant coefficient second order equations is the one where $b^2 - 4ac < 0$. This results in two roots r_1 and r_2, but they are complex roots. In order to solve differential equations with $b^2 - 4ac < 0$, we need to be able to manipulate and use some properties of complex numbers. Complex numbers may seem a strange concept, especially because of the terminology. There is nothing imaginary or really complicated about complex numbers. For more bachground information on complex numbers, see **Normally a reference to an Appendix goes here.**.

To start with, we define $i = \sqrt{-1}$. Since this is the square root of a negative number, this i is not a real number. A complex number is written in the form $z = x + iy$ where x and y are real numbers. For a complex number $x + iy$ we call x the *real part* and y the *imaginary part* of the number. Often the following notation is used,

$$\operatorname{Re}(x+iy) = x \qquad \text{and} \qquad \operatorname{Im}(x+iy) = y.$$

The real numbers are contained in the complex numbers as those complex numbers with the imaginary part being zero.

When trying to do arithmetic with complex numbers, we treat i as though it is a variable, and do computations just as we would with polynomials. The important fact that we will use to simplify is the fact that since $i = \sqrt{-1}$, we have that $i^2 = -1$. So whenever we see i^2, we replace it by -1. For example,

$$(2+3i)(4i) - 5i = (2 \times 4)i + (3 \times 4)i^2 - 5i = 8i + 12(-1) - 5i = -12 + 3i.$$

The numbers i and $-i$ are the two roots of $r^2 + 1 = 0$. Engineers often use the letter j instead of i for the square root of -1. We use the mathematicians' convention and use i.

Exercise 257 *Make sure you understand (that you can justify) the following identities:*

a) $i^2 = -1$, $i^3 = -i$, $i^4 = 1$,

b) $\dfrac{1}{i} = -i$,

c) $(3-7i)(-2-9i) = \cdots = -69 - 13i$,

d) $(3-2i)(3+2i) = 3^2 - (2i)^2 = 3^2 + 2^2 = 13$,

e) $\dfrac{1}{3-2i} = \dfrac{1}{3-2i}\dfrac{3+2i}{3+2i} = \dfrac{3+2i}{13} = \dfrac{3}{13} + \dfrac{2}{13}i.$

In order to solve differential equations where the characteristic equation has complex roots, we need to deal with the exponential e^{a+bi} of complex numbers. We do this by writing down the Taylor series and plugging in the complex number. Because most properties of the exponential can be proved by looking at the Taylor series, these properties still hold for the complex exponential. For example the very important property: $e^{x+y} = e^x e^y$. This means that $e^{a+ib} = e^a e^{ib}$. Hence if we can compute e^{ib}, we can compute e^{a+ib}. For e^{ib}, we use the so-called *Euler's formula.*

Theorem 7 (eulersformula)**.** *Euler's formula*

$$e^{i\theta} = \cos\theta + i\sin\theta \qquad \textit{and} \qquad e^{-i\theta} = \cos\theta - i\sin\theta.$$

In other words, $e^{a+ib} = e^a\big(\cos(b) + i\sin(b)\big) = e^a\cos(b) + ie^a\sin(b)$.

Exercise 258 *Using Euler's formula, check the identities:*

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \textit{and} \qquad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

Learning outcomes: Understand the basics of complex numbers Use complex numbers to find complex solutions to second order constant coefficient equations Use Euler's formula to find real-valued general solutions to these second order equations.

Author(s): Matthew Charnley and Jason Nowell

Exercise 259 *Double angle identities: Start with $e^{i(2\theta)} = \left(e^{i\theta}\right)^2$. Use Euler on each side and deduce:*

$$\cos(2\theta) = \cos^2\theta - \sin^2\theta \qquad \text{and} \qquad \sin(2\theta) = 2\sin\theta\cos\theta.$$

Complex roots

Suppose the equation $ay'' + by' + cy = 0$ has the characteristic equation $ar^2 + br + c = 0$ that has complex roots. By the quadratic formula, the roots are $\dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$. These roots are complex if $b^2 - 4ac < 0$. In this case the roots are

$$r_1, r_2 = \frac{-b}{2a} \pm i\frac{\sqrt{4ac - b^2}}{2a}.$$

As you can see, we always get a pair of roots of the form $\alpha \pm i\beta$. In this case we can still write the solution as

$$y = C_1 e^{(\alpha+i\beta)x} + C_2 e^{(\alpha-i\beta)x}.$$

However, the exponential is now complex-valued, and so (real) linear combinations of these solutions will be complex valued. If we are using these equations to model physical problems, the answer should be real-valued, as the position of a mass-on-a-spring can not be a complex number. To do this, we need to determine two real-valued, linearly independent solutions to this differential equation.

To do this, we use the following result.

Theorem 8 (thm:realimagparts)**.** *Consider the differential equation*

$$y'' + p(x)y' + q(x)y = 0$$

where $p(t)$ and $q(t)$ are real-valued *continuous functions on some interval I. If y is a complex-valued solution to this differential equation and we can split $y(x) = u(x) + iv(x)$ into its real and imaginary parts u and v, then u and v are both solutions to $y'' + p(x)y' + q(x)y = 0$.*

Proof This is based on the fact that the differential equation is linear. We can compute derivatives of y

$$\begin{aligned} y(x) &= u(x) + iv(x) \\ y'(x) &= u'(x) + iv'(x) \\ y''(x) &= u''(x) + iv''(x) \end{aligned}.$$

Then, we can plug this into the differential equation

$$\begin{aligned} 0 &= y'' + p(x)y' + q(x)y \\ &= u''(x) + iv''(x) + p(x)(u'(x) + iv'(x)) + q(x)(u(x) + iv(x)) \\ 0 &= u''(x) + p(x)u'(x) + q(x)u(x) + i(v''(x) + p(x)v'(x) + q(x)v(x)) \end{aligned}.$$

Since the equation at the end of this chain is equal to zero, it must be zero as a complex number, which means that both the real and imaginary parts must be zero. This means that

$$\begin{aligned} u''(x) + p(x)u'(x) + q(x)u(x) &= 0 \\ v''(x) + p(x)v'(x) + q(x)v(x) &= 0 \end{aligned},$$

so that both u and v solve the original differential equation. ■

To use this to solve the problem at hand, we have our solution

$$y_1(x) = e^{\alpha + i\beta x}$$

and we need to split this into its real and imaginary parts. Since

$$y_1 = e^{\alpha x}\cos(\beta x) + ie^{\alpha x}\sin(\beta x),$$

the real and imaginary parts of this function are

$$u(x) = e^{\alpha x} \cos(\beta x)$$
$$v(x) = e^{\alpha x} \sin(\beta x)$$

which, by the previous theorem, we know are also solutions. These are two solutions to our original differential equation that are also real-valued!

On the other hand, assume that we take the other complex solution, which will be

$$y_2(x) = e^{\alpha - i\beta x}.$$

If we split this into real and imaginary parts, we will get

$$y_2 = e^{\alpha x} \cos(\beta x) - ie^{\alpha x} \sin(\beta x),$$

so that the real and imaginary parts of this solution are

$$u_2(x) = e^{\alpha x} \cos(\beta x)$$
$$v_2(x) = -e^{\alpha x} \sin(\beta x).$$

These are exactly the same as the previous real and imaginary parts, up to the minus sign on v_2. Since we are going to incorporate these with constants C_1 and C_2 eventually, they will give rise to the same general solution. So, we only need one of these two complex solutions to generate our two linearly independent real-valued solutions, and either of the two complex solutions give the same pair of real-valued solutions.

Exercise 260 *For $\beta \neq 0$, check that $e^{\alpha x} \cos(\beta x)$ and $e^{\alpha x} \sin(\beta x)$ are linearly independent.*

With that fact, we have the following theorem.

Theorem 9. *Take the equation*

$$ay'' + by' + cy = 0.$$

If the characteristic equation has the roots $\alpha \pm i\beta$ (when $b^2 - 4ac < 0$), then the general solution is

$$y = C_1 e^{\alpha x} \cos(\beta x) + C_2 e^{\alpha x} \sin(\beta x).$$

Example 49. *Find the general solution of $y'' + k^2 y = 0$, for a constant $k > 0$.*

Solution: The characteristic equation is $r^2 + k^2 = 0$. Therefore, the roots are $r = \pm ik$, and by the theorem, we have the general solution

$$y = C_1 \cos(kx) + C_2 \sin(kx).$$

Example 50. *Find the solution of $y'' - 6y' + 13y = 0$, $y(0) = 0$, $y'(0) = 10$.*

Solution: The characteristic equation is $r^2 - 6r + 13 = 0$. By completing the square we get $(r-3)^2 + 2^2 = 0$ and hence the roots are $r = 3 \pm 2i$. By the theorem we have the general solution

$$y = C_1 e^{3x} \cos(2x) + C_2 e^{3x} \sin(2x).$$

To find the solution satisfying the initial conditions, we first plug in zero to get

$$0 = y(0) = C_1 e^0 \cos 0 + C_2 e^0 \sin 0 = C_1.$$

Hence, $C_1 = 0$ and $y = C_2 e^{3x} \sin(2x)$. We differentiate,

$$y' = 3C_2 e^{3x} \sin(2x) + 2C_2 e^{3x} \cos(2x).$$

We again plug in the initial condition and obtain $10 = y'(0) = 2C_2$, or $C_2 = 5$. The solution we are seeking is

$$y = 5e^{3x} \sin(2x).$$

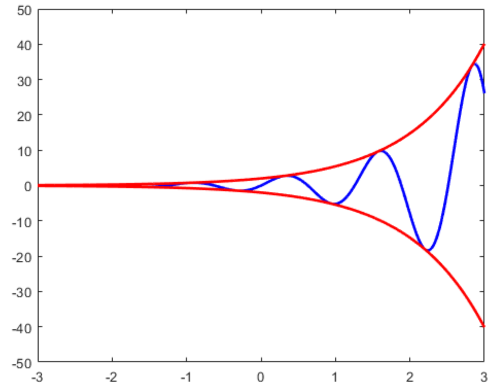


Figure 38: Plot of the function $y = 2e^x \sin(5x)$ with envelope curves.

In this previous example, we can get a fairly good idea of how to sketch out the graph of this function. Since $\sin(2x)$ oscillates between -1 and 1, the graph of $y = 5e^{3x}\sin(2x)$ will oscillate between the graphs of $5e^{3x}$ and $-5e^{3x}$. These curves that surround the graph of the solution are called *envelope curves* for the solution. In Figure 38, this phenomenon is illustrated for the function $y = 2e^x \sin(5x)$.

This is simple when there is only one term in the function we want to draw. When both sine and cosine terms appear, this can get more tricky, but we can still work it out. In the more general case, the solution will look something like

$$y = Ae^{\alpha x}\cos(\beta x) + Be^{\alpha x}\sin(\beta x).$$

We can first factor out an $e^{\alpha x}$, and then we want to write $A\cos(\beta x) + B\sin(\beta x)$ as a single trigonometric function. The identity we want to use here is the trigonometric identity

$$\cos(\beta x - \delta) = \cos(\delta)\cos(\beta x) + \sin(\delta)\sin(\beta x).$$

If there is an angle δ so that $A = \cos(\delta)$ and $B = \sin(\delta)$, then we could write

$$A\cos(\beta x) + B\sin(\beta x) = \cos(\beta x - \delta)$$

and we would be done. However, this does not always happen; the main issue being that $\cos^2(\delta) + \sin^2(\delta) = 1$ for all δ, but it is not necessarily the case that $A^2 + B^2 = 1$. But we can force this last condition. If we define $R = \sqrt{A^2 + B^2}$, then we can rewrite this expression as

$$\begin{aligned} A\cos(\beta x) + B\sin(\beta x) &= R\left(\frac{A}{\sqrt{A^2+B^2}}\cos(\beta x) + \frac{B}{\sqrt{A^2+B^2}}\sin(\beta x)\right) \\ &= R\left(\cos(\delta)\cos(\beta x) + \sin(\delta)\sin(\beta x)\right) \\ &= R\cos(\beta x - \delta) \end{aligned}$$

where δ is the angle so that

$$\cos(\delta) = \frac{A}{R} \qquad \sin(\delta) = \frac{B}{R}$$

and such an angle will always exist. Therefore, we can represent the original solution

$$y = Ae^{\alpha x}\cos(\beta x) + Be^{\alpha x}\sin(\beta x)$$

as

$$y = Re^{\alpha x}\cos(\beta x - \delta)$$

where

$$R = \sqrt{A^2 + B^2} \qquad \cos(\delta) = \frac{A}{R} \qquad \sin(\delta) = \frac{B}{R}.$$

Therefore, the envelope curves for this solution will be

$$y = \pm R e^{\alpha x}.$$

Note that in order to determine these envelope curves, you do not need to determine the δ value in the representation of the solution. All you need is the value of R, which can be computed as $\sqrt{A^2 + B^2}$ where A and B are the coefficients of the sine and cosine terms in the solution.

Example 51. *Find the solution to the initial value problem*

$$y'' + 2y' + 5y = 0 \qquad y(0) = 1,\ y'(0) = 5.$$

Determine a value T where the solution $y(x)$ satisfies $|y(x)| < 0.1$ for all $x > T$.

Solution: We solve the initial value problem by normal techniques from this section. The characteristic equation is $r^2 + 2r + 5 = 0$, which has roots $r = -1 \pm 2i$. Therefore, the general solution of the differential equation is

$$y = C_1 e^{-x}\cos(2x) + C_2 e^{-x}\sin(2x).$$

Plugging in 0 gives that $y(0) = 1 = C_1$, and the derivative of this general solution is

$$y' = -C_1 e^{-x}\cos(2x) - 2C_1 e^{-x}\sin(2x) - C_2 e^{-x}\sin(2x) + 2C_2 e^{-x}\cos(2x).$$

Plugging in 0 here gives

$$y'(0) = -C_1 + 2C_2.$$

Since $C_1 = 1$, this gives that $C_2 = 3$. So, our solution is

$$y(x) = e^{-x}\cos(2x) + 3e^{-x}\sin(2x).$$

Through the work above, we can find $R = \sqrt{1+9} = \sqrt{10}$. Therefore, the envelope curves for the solution are

$$\pm\sqrt{10}e^{-x}.$$

In order to find this threshold T where the solution will stay within 0.1 of zero, we need to figure out when this envelope curves get to the 0.1 threshold. Once the envelope curves get to that level, we know that the full solution must be trapped there as well. We can solve

$$0.1 = \sqrt{10}e^{-T} \qquad T = -\ln\left(\frac{0.1}{\sqrt{10}}\right) \approx 3.454.$$

So, for all values of x larger than 3.454, the solution will be within 0.1 of zero. This is illustrated in Figure 39. Note that we did not find the *best* value T here, as it probably could be made smaller using the actual solution. The issue here is that because the solution is oscillating, it may end up staying inside the 0.1 cutoff before that value of time, but this is the lowest value of T that we can prove and validate using envelope curves.

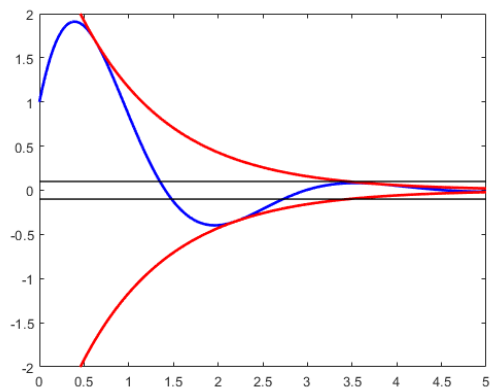


Figure 39: Plot of the function $e^{-x}\cos(2x) + 3e^{-x}\sin(2x)$ with envelope curves illustrating the bounds on the function for large values of x.

higherOrder/practice/complexRoots-practice1.tex

Practice for Complex Roots

Why?

Exercise 261 *Write* $3\cos(2x) + 3\sin(2x)$ *in the form* $R\cos(\beta x - \delta)$.

Exercise 262 *Write* $2\cos(3x) + \sin(3x)$ *in the form* $R\cos(\beta x - \delta)$.

Exercise 263 *Write* $3\cos(x) - 4\sin(x)$ *in the form* $R\cos(\beta x - \delta)$.

Exercise 264 *Show that* $e^{2x}\cos(x)$ *and* $e^{2x}\sin(x)$ *are linearly independent.*

Exercise 265 *Find the general solution of* $2y'' + 50y = 0$.

Exercise 266 *Find the general solution of* $y'' - 6y' + 13y = 0$.

Exercise 267 *Find the solution to* $y'' - 2y' + 5y = 0$ *with* $y(0) = 3$ *and* $y'(0) = 2$.

Exercise 268 *Find the general solution of* $y'' + 2y' - 3y = 0$.

Exercise 269 *Find the solution to* $2y'' + y' + y = 0$, $y(0) = 1$, $y'(0) = -2$.

Exercise 270 *Find the solution to* $z''(t) = -2z'(t) - 2z(t)$, $z(0) = 2$, $z'(0) = -2$.

Exercise 271 *Let us revisit the Cauchy–Euler equations of* ***Normally a reference to a previous exercise goes here.****. Suppose now that* $(b-a)^2 - 4ac < 0$. *Find a formula for the general solution of* $ax^2y'' + bxy' + cy = 0$. *Hint: Note that* $x^r = e^{r\ln x}$.

Exercise 272 *Construct an equation such that* $y = C_1e^{-2x}\cos(3x) + C_2e^{-2x}\sin(3x)$ *is the general solution.*

Author(s): Matthew Charnley and Jason Nowell

Exercise 273 *Find a second order, constant coefficient differential equation with general solution given by* $y(t) = C_1 e^x \cos(2x) + C_2 e^{2x} \sin(x)$ *or explain why there is no such thing.*

Exercise 274 *Find a second order, constant coefficient differential equation with general solution given by* $y(t) = C_1 e^x \cos(2x) + C_2 e^x \sin(2x)$ *or explain why there is no such thing.*

Exercise 275 *Find the solution to the initial value problem*

$$y'' + 4y' + 5y = 0 \qquad y(0) = 3,\ y'(0) = -1.$$

Determine a value T *so that* $|y(x)| < 0.02$ *for all* $x > T$.

Exercise 276 *Find the solution to the initial value problem*

$$y'' + 6y' + 13y = 0 \qquad y(0) = 4,\ y'(0) = 7.$$

Determine a value T *so that* $|y(x)| < 0.01$ *for all* $x > T$.

Repeated Roots and Reduction of Order

We discuss Repeated Roots and Reduction of Order

The last case we have to handle for solving all second order linear constant coefficient equations is the case where $b^2-4ac=0$ in the equation

$$ay''+by'+cy=0.$$

When we try to find the characteristic equation and find solutions to this equation, we get a double root at r_1, so that the characteristic polynomial is $(r-r_1)^2$. For this, we get that e^{r_1x} is a solution. However, that's the only solution we get. We need to have two linearly independent solutions in order to get the general solution to the differential equation, so we need to find some method to get another solution. The standard method, and the one we apply here is *reduction of order*. Let's see how this works through an example.

Example 52. *Find two linearly independent solutions to the differential equation*

$$y''+2y'+y=0.$$

Solution: To start, we find the first solution using our original method. The characteristic equation here is $r^2+2r+1=0$, which is $(r+1)^2$. Therefore, we have a double root at $r=-1$, so that $y_1(x)=e^{-x}$ is a solution.

To find a second solution, the reduction of order method suggests that we try to plug in $y=v(x)e^{-x}$ for an unknown function $v(x)$. The goal is to figure out an equation that v must satisfy to see if this leads us to a second solution to the original equation. We can compute the first two derivatives of $y=v(x)e^{-x}$

$$\begin{aligned} y(x) &= v(x)e^{-x} \\ y'(x) &= v'(x)e^{-x}-v(x)e^{-x} \\ y''(x) &= v''(x)e^{-x}-2v'(x)e^{-x}+v(x)e^{-x} \end{aligned}$$

and then plug them into the original differential equation

$$\begin{aligned} 0 &= y''+2y'+y \\ &= (v''(x)e^{-x}-2v'(x)e^{-x}+v(x)e^{-x})+2(v'(x)e^{-x}-v(x)e^{-x})+v(x)e^{-x} \\ &= v''(x)e^{-x}+v'(x)(-2e^{-x}+2e^{-x})+v(x)(e^{-x}-2e^{-x}+e^{-x}) \\ &= v''(x)e^{-x} \end{aligned}.$$

Since e^{-x} is never zero, this means we must have $v''(x)=0$. This is still a second order equation, but we know how to solve it. We can integrate both sides twice to get that $v(x)=Ax+B$ for any constants A and B.

Our goal with all of this was to find a solution y of the form $v(x)e^{-x}$. The set up here means that $y=(Ax+B)e^{-x}$ will solve the differential equation. Since we already knew that Be^{-x} was a solution, the new information we gained here was that Axe^{-x}, or in particular, xe^{-x} is a solution to the differential equation. Thus, our two solutions are $y_1(x)=e^{-x}$ and $y_2(x)=xe^{-x}$.

Exercise 277 *Check that e^{-x} and xe^{-x} both solve $y''+2y'+y=0$, and that these solutions are linearly independent.*

The *reduction of order method* applies more generally to any second order linear homogeneous equation and the goal is the same: use one solution of the differential equation to generate another one. The idea is that if we somehow found y_1 as a solution of $y''+p(x)y'+q(x)y=0$ we try a second solution of the form $y_2(x)=y_1(x)v(x)$. We just need to find v. We plug y_2 into the equation:

$$\begin{aligned} 0=y_2''+p(x)y_2'+q(x)y_2 &= y_1''v+2y_1'v'+y_1v''+p(x)(y_1'v+y_1v')+q(z)y_1v \\ &= y_1v''+(2y_1'+p(x)y_1)v'+\cancelto{0}{\left(y_1''+p(x)y_1'+q(x)y_1\right)}\quad v. \end{aligned}$$

Learning outcomes: Find the general solution to a second order constant coefficient equation with repeated roots Apply the method of reduction of order to generate a second solution to an equation given one solution Solve Euler equations using the method of reduction of order.

Author(s): Matthew Charnley and Jason Nowell

In other words, $y_1 v'' + (2y_1' + p(x)y_1)v' = 0$. Using $w = v'$ we have the first order linear equation $y_1 w' + (2y_1' + p(x)y_1)w = 0$. After solving this equation for w (integrating factor), we find v by antidifferentiating w. We then form y_2 by computing $y_1 v$. For example, suppose we somehow know $y_1 = x$ is a solution to $y'' + x^{-1}y' - x^{-2}y = 0$. The equation for w is then $xw' + 3w = 0$. We find a solution, $w = Cx^{-3}$, and we find an antiderivative $v = \frac{-C}{2x^2}$. Hence $y_2 = y_1 v = \frac{-C}{2x}$. Any C works and so $C = -2$ makes $y_2 = 1/x$. Thus, the general solution is $y = C_1 x + C_2 1/x$.

The easiest way to work out these problems is to remember that we need to try $y_2(x) = y_1(x)v(x)$ and find $v(x)$ as we did above. Also, the technique works for higher order equations too: you get to reduce the order for each solution you find.

In summary, for constant coefficient equations with a repeated root, the reduction of order method will always give the equation $v'' = 0$, and so the solution is $v(x) = Ax + B$. Multiplying by the y_1 solution e^{rx} gives that xe^{rx} is the other solution. Therefore, the general solution for repeated root equations is always of the form

$$y = C_1 e^{r_1 x} + C_2 x r^{r_1 x}.$$

Example 53. *Find the general solution of*

$$y'' - 8y' + 16y = 0.$$

Solution: The characteristic equation is $r^2 - 8r + 16 = (r-4)^2 = 0$. The equation has a double root $r_1 = r_2 = 4$. The general solution is, therefore,

$$y = (C_1 + C_2 x)\, e^{4x} = C_1 e^{4x} + C_2 x e^{4x}.$$

Exercise 278 *Check that e^{4x} and xe^{4x} are linearly independent.*

That e^{4x} solves the equation is clear. If xe^{4x} solves the equation, then we know we are done. Let us compute $y' = e^{4x} + 4xe^{4x}$ and $y'' = 8e^{4x} + 16xe^{4x}$. Plug in

$$y'' - 8y' + 16y = 8e^{4x} + 16xe^{4x} - 8(e^{4x} + 4xe^{4x}) + 16xe^{4x} = 0.$$

In some sense, a doubled root rarely happens. If coefficients are picked randomly, a doubled root is unlikely. There are, however, some natural phenomena where a doubled root does happen, so we cannot just dismiss this case. In addition, there are specific physical applications that involve the double root problem, which we will discuss in Section . Finally, the solution with a doubled root can be thought of as an approximation of the solution with two roots that are very close together, and the behavior of this solution will approximate "nearby" solutions as well.

Example 54. *Find the solution $y(t)$ to the initial value problem*

$$y'' + 6y' + 9y = 0 \qquad y(0) = 2,\ y'(0) = -3.$$

Solution: The characteristic polynomials for this differential equation is

$$r^2 + 6r + 9$$

which factors as $(r+3)^2$, so that we have a double root at -3. With the work done previously, we know that the general solution is

$$y(t) = (C_1 + C_2 t)e^{-3t} = C_1 e^{-3t} + C_2 t e^{-3t}.$$

If we use the initial conditions, we can set $t = 0$ to get that

$$2 = y(0) = C_1 e^0$$

so that $C_1 = 2$. Differentiating the general solution gives that

$$y'(t) = -3C_1 e^{-3t} + C_2 e^{-3t} - 3C_2 t e^{-3t}$$

and plugging in zero here gives

$$-3 = y'(0) = -3C_1 + C_2.$$

Since $C_1 = 2$, this implies that $C_2 = 3$. Therefore, the solution to this initial value problem is

$$y(t) = 2e^{-3t} + 3te^{-3t}.$$

`higherOrder/practice/repeatedRoots-practice1.tex`

Practice for Second Order Repeated Roots

Why?

Exercise 279 *Find the general solution to* $y'' + 4y' + 4y = 0$.

Exercise 280 *Find the general solution to* $y'' - 6y' + 9y = 0$.

Exercise 281 *Find the solution to* $y'' + 6y' + 9y = 0$ *with* $y(0) = 3$ *and* $y'(0) = -1$.

Exercise 282 *Solve* $y'' - 8y' + 16y = 0$ *for* $y(0) = 2$, $y'(0) = 0$.

Exercise 283 *Find the general solution of* $y'' = 0$ *using the methods of this section.*

Exercise 284 *The method of this section applies to equations of other orders than two. We will see higher orders later. Try to solve the first order equation* $2y' + 3y = 0$ *using the methods of this section.*

Exercise 285 *Consider the second-order DE*

$$ty'' + (4t + 2)y' + (4t + 4)y = 0. \tag{21}$$

a) *Does the* superposition principle *apply to this DE? Give a one- or two-sentence explanation wither way.*

b) *Find a value of* r *so that* $y = e^{rt}$ *is a solution to* (21)

c) *Using your result from the previous page, apply* **reduction of order** *to find the general solution to* (21).

Exercise 286 *Consider the differential equation* $x^2y'' + 3xy' - 3y = 0$.

a) *Verify that* $y_1(x) = x$ *is a solution.*

b) *Use reduction of order to find a second linearly independent solution.*

c) *Write out the general solution.*

Exercise 287 *Consider the differential equation* $x^2y'' + 4xy' + 2y = 0$.

Author(s): Matthew Charnley and Jason Nowell

a) *Verify that* $y_1(x) = \frac{1}{x}$ *is a solution.*

b) *Use reduction of order to find a second linearly independent solution.*

c) *Write out the general solution.*

Exercise 288 *Consider the differential equation* $x^2y'' - 6xy' + 10y = 0$.

a) *Verify that* $y_1(x) = x^2$ *is a solution.*

b) *Use reduction of order to find a second linearly independent solution.*

c) *Write out the general solution.*

Exercise 289 *Write down a differential equation with general solution* $y = at^2 + bt^{-3}$, *or explain why there is no such thing.*

Exercise 290 *Find the solution to* $y'' - (2\alpha)y' + \alpha^2 y = 0$, $y(0) = a$, $y'(0) = b$, *where* α, a, *and* b *are real numbers.*

Exercise 291 *Suppose* y_1 *is a solution to* $y'' + p(x)y' + q(x)y = 0$. *By directly plugging into the equation, show that*

$$y_2(x) = y_1(x) \int \frac{e^{-\int p(x)\,dx}}{\big(y_1(x)\big)^2}\,dx$$

is also a solution.

Exercise 292 *Take* $(1 - x^2)y'' - xy' + y = 0$.

a) *Show that* $y = x$ *is a solution.*

b) *Use reduction of order to find a second linearly independent solution.*

c) *Write down the general solution.*

Exercise 293 *Take* $y'' - 2xy' + 4y = 0$.

a) *Show that* $y = 1 - 2x^2$ *is a solution.*

b) *Use reduction of order to find a second linearly independent solution. (It's OK to leave a definite integral in the formula.)*

c) *Write down the general solution.*

The rest of these exercises can be solved using any of the methods discussed in the last three sections. Pick the appropriate method in order to solve the problem.

Exercise 294 *Find the general solution of* $y'' + 5y' - 6y = 0$.

Exercise 295 *Find the general solution of* $y'' - 2y' + 2y = 0$.

Exercise 296 *Find the general solution of* $y'' + 4y' + 4y = 0$.

Exercise 297 *Find the general solution of* $y'' + 4y' + 5y = 0$.

Exercise 298 *Find the solution to* $y'' - 6y' + 13y = 0$ *with* $y(0) = 2$ *and* $y'(0) = 1$.

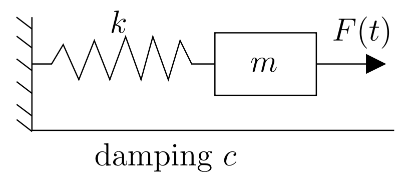
Exercise 299 *Find the solution to* $y'' + 4y' - 12y = 0$ *with* $y(0) = -1$ *and* $y'(0) = 3$.

Exercise 300 *Find the solution to* $y'' - 6y' + 9y = 0$ *with* $y(0) = -4$ *and* $y'(0) = -1$.

Mechanical vibrations

We discuss Mechanical vibrations

In the last few sections, we have discussed all of the different possible solutions to constant coefficient second order differential equations, whether the roots of the characteristic polynomial real and distinct, complex, or repeated. Now, we want to look at applications of these equations, now that we know how to solve them. Since Newton's Second Law $F = ma$ involves the second derivative of position (acceleration), it is reasonable that a lot of physical systems will be defined by second order differential equations.



Our first example is a mass on a spring. Suppose we have a mass $m > 0$ (in kilograms) connected by a spring with spring constant $k > 0$ (in newtons per meter) to a fixed wall. There may be some external force $F(t)$ (in newtons) acting on the mass. Finally, there is some friction measured by $c \geq 0$ (in newton-seconds per meter) as the mass slides along the floor (or perhaps a damper is connected).

Let x be the displacement of the mass ($x = 0$ is the rest position), with x growing to the right (away from the wall). The force exerted by the spring is proportional to the compression of the spring by Hooke's law. Therefore, it is kx in the negative direction. Similarly the amount of force exerted by friction is proportional to the velocity of the mass. By Newton's second law we know that force equals mass times acceleration and hence $mx'' = F(t) - cx' - kx$ or

$$mx'' + cx' + kx = F(t).$$

This is a linear second order constant coefficient ODE. We say the motion is

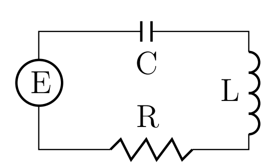
(a) *forced*, if $F \not\equiv 0$ (if F is not identically zero),

(b) *unforced* or *free*, if $F \equiv 0$ (if F is identically zero),

(c) *damped*, if $c > 0$, and

(d) *undamped*, if $c = 0$.

This system appears in lots of applications even if it does not at first seem like it. Many real-world scenarios can be simplified to a mass on a spring. For example, a bungee jump setup is essentially a mass and spring system (you are the mass). It would be good if someone did the math before you jump off the bridge, right? Let us give two other examples.



Here is an example for electrical engineers. Consider the pictured RLC circuit. There is a resistor with a resistance of R ohms, an inductor with an inductance of L henries, and a capacitor with a capacitance of C farads. There is also an electric source (such as a battery) giving a voltage of $E(t)$ volts at time t (measured in seconds). Let $Q(t)$ be the charge in coulombs on the capacitor and $I(t)$ be the current in the circuit. The relation between the two is$Q' = I$. By elementary principles we find $LI' + RI + Q/C = E$. Since $Q' = I$, this means that $I' = Q''$, and we can write this equation as

$$LQ''(t) + RQ'(t) + \frac{1}{C}Q(t) = E(t).$$

We can also write this a different way by differentiating the entire equation in t to get a second order equation for $I(t)$:

$$LI''(t) + RI'(t) + \frac{1}{C}I(t) = E'(t).$$

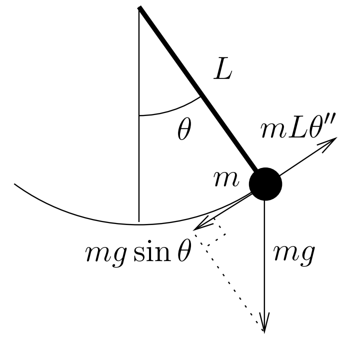
This is a nonhomogeneous second order constant coefficient linear equation. As L, R, and C are all positive, this system behaves just like the mass and spring system. Position of the mass is replaced by current. Mass is replaced by inductance, damping is replaced by resistance, and the spring constant is replaced by one over the capacitance. The change in voltage becomes the forcing function—for constant voltage this is an unforced motion.

Learning outcomes: Write second-order differential equations to model physical situations Classify a mechanical oscillation as undamped, underdamped, critically damped, or overdamped Use the solution to a differential equation to describe the resulting physical motion.

Author(s): Matthew Charnley and Jason Nowell

Our next example behaves like a mass and spring system only approximately. Suppose a mass m hangs on a pendulum of length L. We seek an equation for the angle $\theta(t)$ (in radians). Let g be the force of gravity. Elementary physics mandates that the equation is

$$\theta'' + \frac{g}{L}\sin\theta = 0.$$



Let us derive this equation using Newton's second law: force equals mass times acceleration. The acceleration is $L\theta''$ and mass is m. So $mL\theta''$ has to be equal to the tangential component of the force given by the gravity, which is $mg\sin\theta$ in the opposite direction. So $mL\theta'' = -mg\sin\theta$. The m curiously cancels from the equation.

Now we make our approximation. For small θ we have that approximately $\sin\theta \approx \theta$. This can be seen by looking at the graph. In **Normally a reference to a previous figure goes here.** we can see that for approximately $-0.5 < \theta < 0.5$ (in radians) the graphs of $\sin\theta$ and θ are almost the same.

Therefore, when the swings are small, θ is small and we can model the behavior by the simpler linear equation

$$\theta'' + \frac{g}{L}\theta = 0.$$

The errors from this approximation build up. So after a long time, the state of the real-world system might be substantially different from our solution. Also we will see that in a mass-spring system, the amplitude is independent of the period. This is not true for a pendulum. Nevertheless, for reasonably short periods of time and small swings (that is, only small angles θ), the approximation is reasonably good.

In real-world problems it is often necessary to make these types of simplifications. We must understand both the mathematics and the physics of the situation to see if the simplification is valid in the context of the questions we are trying to answer.

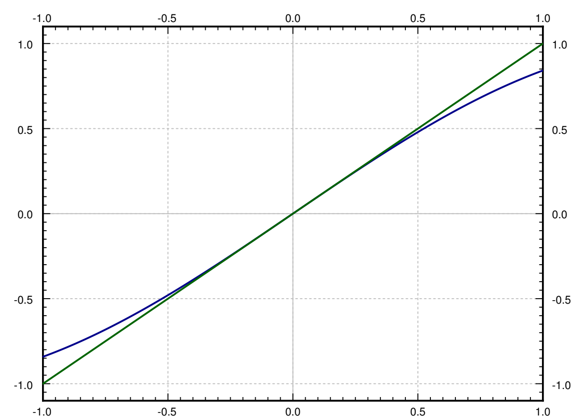


Figure 40: The graphs of $\sin\theta$ and θ (in radians).

Free undamped motion

In this section we only consider free or unforced motion, as we do not know yet how to solve nonhomogeneous equations. Let us start with undamped motion where $c = 0$. The equation is

$$mx'' + kx = 0.$$

We divide by m and let $\omega_0 = \sqrt{k/m}$ to rewrite the equation as

$$x'' + \omega_0^2 x = 0.$$

The general solution to this equation is

$$x(t) = A\cos(\omega_0 t) + B\sin(\omega_0 t).$$

By a trigonometric identity that we discussed previously in § ,

$$A\cos(\omega_0 t) + B\sin(\omega_0 t) = C\cos(\omega_0 t - \delta),$$

for two constants C and γ. Earlier, we found that we can compute these constants as $C = \sqrt{A^2 + B^2}$ and $\tan\delta = B/A$. Therefore, we let C and δ be our arbitrary constants and write $x(t) = C\cos(\omega_0 t - \delta)$.

Exercise 301 *Justify the identity $A\cos(\omega_0 t) + B\sin(\omega_0 t) = C\cos(\omega_0 t - \delta)$ and verify the equations for C and δ. Hint: Start with $\cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)$ and multiply by C. Then what should α and β be?*

While it is generally easier to use the first form with A and B to solve for the initial conditions, the second form is much more natural to use for interpretation of physical systems, since the constants C and δ have nice physical interpretation. Write the solution as

$$x(t) = C\cos(\omega_0 t - \delta).$$

This is a pure-frequency oscillation (a sine wave). The *amplitude* is C, ω_0 is the (angular) *frequency*, and δ is the so-called *phase shift*. The phase shift just shifts the graph left or right. We call ω_0 the *natural (angular) frequency*. This entire setup is called *simple harmonic motion*.

Let us pause to explain the word *angular* before the word *frequency*. The units of ω_0 are radians per unit time, not cycles per unit time as is the usual measure of frequency. Because one cycle is 2π radians, the usual frequency is given by $\frac{\omega_0}{2\pi}$. It is simply a matter of where we put the constant 2π, and that is a matter of taste.

The *period* of the motion is one over the frequency (in cycles per unit time) and hence $\frac{2\pi}{\omega_0}$. That is the amount of time it takes to complete one full cycle.

Example 55. *Suppose that $m = 2\,\mathrm{kg}$ and $k = 8\,{}^{\mathrm{N}}\!/\!{}_{\mathrm{m}}$. The whole mass and spring setup is sitting on a truck that was traveling at $1\,{}^{m}\!/\!{}_{s}$. The truck crashes and hence stops. The mass was held in place 0.5 meters forward from the rest position. During the crash the mass gets loose. That is, the mass is now moving forward at $1\,{}^{m}\!/\!{}_{s}$, while the other end of the spring is held in place. The mass therefore starts oscillating. What is the frequency of the resulting oscillation? What is the amplitude? The units are the mks units (meters-kilograms-seconds).*

Solution: The setup means that the mass was at half a meter in the positive direction during the crash and relative to the wall the spring is mounted to, the mass was moving forward (in the positive direction) at $1\,{}^{\mathrm{m}}\!/\!{}_{\mathrm{s}}$. This gives us the initial conditions.

So the equation with initial conditions is

$$2x'' + 8x = 0, \qquad x(0) = 0.5, \qquad x'(0) = 1.$$

We directly compute $\omega_0 = \sqrt{{}^{k}\!/\!{}_{m}} = \sqrt{4} = 2$. Hence the angular frequency is 2. The usual frequency in Hertz (cycles per second) is ${}^{2}\!/\!{}_{2\pi} = {}^{1}\!/\!{}_{\pi} \approx 0.318$.

The general solution is

$$x(t) = A\cos(2t) + B\sin(2t).$$

Letting $x(0) = 0.5$ means $A = 0.5$. Then $x'(t) = -2(0.5)\sin(2t) + 2B\cos(2t)$. Letting $x'(0) = 1$ we get $B = 0.5$. Therefore, the amplitude is $C = \sqrt{A^2 + B^2} = \sqrt{0.25 + 0.25} = \sqrt{0.5} \approx 0.707$. The solution is

$$x(t) = 0.5\cos(2t) + 0.5\sin(2t).$$

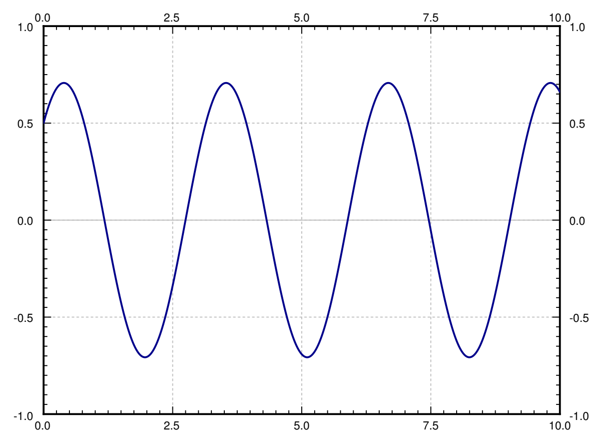


Figure 41: Simple undamped oscillation.

A plot of $x(t)$ is shown in **Normally a reference to a previous figure goes here.**.

In general, for free undamped motion, a solution of the form

$$x(t) = A\cos(\omega_0 t) + B\sin(\omega_0 t),$$

corresponds to the initial conditions $x(0) = A$ and $x'(0) = \omega_0 B$. Therefore, it is easy to figure out A and B from the initial conditions. The amplitude and the phase shift can then be computed from A and B. In the example, we have already found the amplitude C. Let us compute the phase shift. We know that $\tan\delta = {}^{B}\!/\!{}_{A} = 1$. We take the arctangent of 1 and get ${}^{\pi}\!/\!{}_{4}$ or approximately 0.785. We still need to check if this δ is in the correct quadrant (and add π to δ if it is not). Since both A and B are positive, then δ should be in the first quadrant, ${}^{\pi}\!/\!{}_{4}$ radians is in the first quadrant, so $\delta = {}^{\pi}\!/\!{}_{4}$.

Note: Many calculators and computer software have not only the `atan` function for arctangent, but also what is sometimes called `atan2`. This function takes two arguments, B and A, and returns a δ in the correct quadrant for you.

Free damped motion

Let us now focus on damped motion. Let us rewrite the equation

$$mx'' + \gamma x' + kx = 0,$$

as

$$x'' + 2px' + \omega_0^2 x = 0,$$

where

$$\omega_0 = \sqrt{\frac{k}{m}}, \qquad p = \frac{\gamma}{2m}.$$

The characteristic equation is

$$r^2 + 2pr + \omega_0^2 = 0.$$

Using the quadratic formula we get that the roots are

$$r = -p \pm \sqrt{p^2 - \omega_0^2}.$$

The form of the solution depends on whether we get complex or real roots. We get real roots if and only if the following number is nonnegative:

$$p^2 - \omega_0^2 = \left(\frac{\gamma}{2m}\right)^2 - \frac{k}{m} = \frac{\gamma^2 - 4km}{4m^2}.$$

The sign of $p^2 - \omega_0^2$ is the same as the sign of $\gamma^2 - 4km$. Thus we get real roots if and only if $\gamma^2 - 4km$ is nonnegative, or in other words if $\gamma^2 \geq 4km$. If these look familiar, that is not surprising, as they are the same as the conditions we had for the different types of roots in second order constant coefficient equations.

Overdamping

When $\gamma^2 - 4km > 0$, the system is *overdamped*. In this case, there are two distinct real roots r_1 and r_2. Both roots are negative: As $\sqrt{p^2 - \omega_0^2}$ is always less than p, then $-p \pm \sqrt{p^2 - \omega_0^2}$ is negative in either case.

The solution is

$$x(t) = C_1 e^{r_1 t} + C_2 e^{r_2 t}.$$

Since r_1, r_2 are negative, $x(t) \to 0$ as $t \to \infty$. Thus the mass will tend towards the rest position as time goes to infinity. For a few sample plots for different initial conditions, see **Normally a reference to a previous figure goes here.**.

No oscillation happens. In fact, the graph crosses the x-axis at most once. To see why, we try to solve $0 = C_1 e^{r_1 t} + C_2 e^{r_2 t}$. Therefore, $C_1 e^{r_1 t} = -C_2 e^{r_2 t}$ and using laws of exponents we obtain

$$\frac{-C_1}{C_2} = e^{(r_2 - r_1)t}.$$

This equation has at most one solution $t \geq 0$. For some initial conditions the graph never crosses the x-axis, as is evident from the sample graphs.

Example 56. *Suppose the mass is released from rest. That is $x(0) = x_0$ and $x'(0) = 0$. Then*

$$x(t) = \frac{x_0}{r_1 - r_2}\left(r_1 e^{r_2 t} - r_2 e^{r_1 t}\right).$$

It is not hard to see that this satisfies the initial conditions.

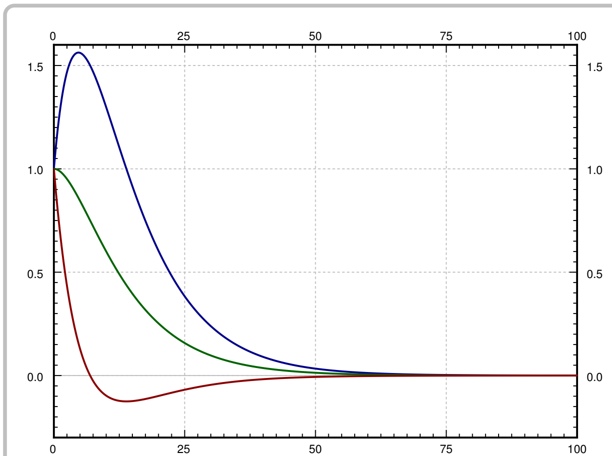


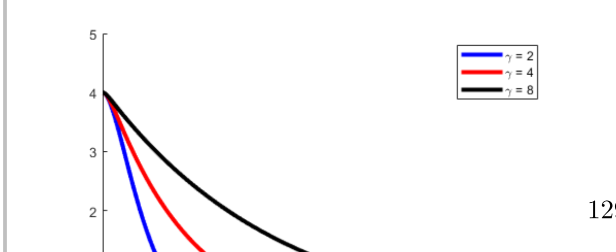
Figure 42: Overdamped motion for several different initial conditions.

Critical damping

When $\gamma^2 - 4km = 0$, the system is *critically damped*. In this case, there is one root of multiplicity 2 and this root is $-p$. Our solution is

$$x(t) = C_1 e^{-pt} + C_2 t e^{-pt}.$$

The behavior of a critically damped system is very similar to an overdamped system. After all a critically damped system is in some sense a limit of overdamped systems. Even though our models are only approximations of the real world problem, the idea of critical damping can be helpful in optimizing systems.



Normally a reference to a previous figure goes here. shows how the solution to

$$x'' + \gamma x' + x = 0$$

for different values of γ and initial conditions $x(0) = 4$ and $x'(0) = 0$. This solution is critically damped if $\gamma = 2$, as that will give us a repeated root in the characteristic equation. Comparing these solutions, we see that the critically damped solution gets back to equilibrium faster than any of the more overdamped solution. When trying to design a system, if we want it to settle back to the zero point as quickly as possible, then we should try to get as closed to critically damped as possible. Even though we are always a little bit underdamped or a little bit overdamped, getting as close as possible will give the best possible result for returning to equilibrium.

Underdamping

When $\gamma^2 - 4km < 0$, the system is *underdamped*. In this case, the roots are complex.

$$\begin{aligned} r &= -p \pm \sqrt{p^2 - \omega_0^2} \\ &= -p \pm \sqrt{-1}\sqrt{\omega_0^2 - p^2} \\ &= -p \pm i\omega_1, \end{aligned}$$

where $\omega_1 = \sqrt{\omega_0^2 - p^2}$. Our solution is

$$x(t) = e^{-pt}\big(A\cos(\omega_1 t) + B\sin(\omega_1 t)\big),$$

or

$$x(t) = Ce^{-pt}\cos(\omega_1 t - \delta).$$

An example plot is given in **Normally a reference to a previous figure goes here.**. Note that we still have that $x(t) \to 0$ as $t \to \infty$.

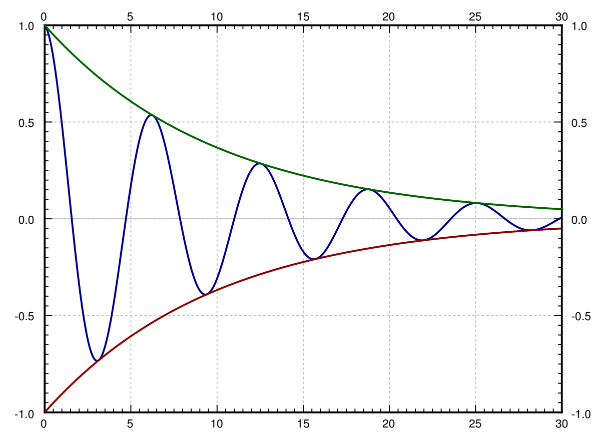


Figure 44: Underdamped motion with the envelope curves shown.

The figure also shows the *envelope curves* Ce^{-pt} and $-Ce^{-pt}$. The solution is the oscillating line between the two envelope curves. The envelope curves give the maximum amplitude of the oscillation at any given point in time. For example, if you are bungee jumping, you are really interested in computing the envelope curve as not to hit the concrete with your head.

The phase shift δ shifts the oscillation left or right, but within the envelope curves (the envelope curves do not change if δ changes).

Notice that the angular *pseudo-frequency*[1] or *quasi-frequency* becomes smaller when the damping γ (and hence p) becomes larger. This makes sense. When we change the damping just a little bit, we do not expect the behavior of the solution to change dramatically. If we keep making γ larger, then at some point the solution should start looking like the solution for critical damping or overdamping, where no oscillation happens. So if γ^2 approaches $4km$, we want ω_1 to approach 0.

Since $\omega_1 = \sqrt{\omega_0^2 - p^2}$ with $p = \dfrac{\gamma}{2m}$ and $\omega_0 = \sqrt{\dfrac{k}{m}}$, we have that

$$\omega_1 = \sqrt{\frac{k}{m} - \frac{\gamma^2}{4m^2}} = \sqrt{\frac{4mk - \gamma^2}{4m^2}},$$

which does go to zero as γ^2 gets closer to $4mk$.

On the other hand, when γ gets smaller, ω_1 approaches ω_0 (ω_1 is always smaller than ω_0), and the solution looks more and more like the steady periodic motion of the undamped case. The envelope curves become flatter and flatter as γ (and hence p) goes to 0.

[1] We do not call ω_1 a frequency since the solution $x(t)$ is not really a periodic function.

`higherOrder/practice/mechanicalVibration-practice1.tex`

Practice for Mechanical Vibrations

Why?

Exercise 302 *Consider a mass and spring system with a mass $m = 2$, spring constant $k = 3$, and damping constant $\gamma = 1$.*

a) *Set up and find the general solution of the system.*

b) *Is the system underdamped, overdamped or critically damped?*

c) *If the system is not critically damped, find a γ that makes the system critically damped.*

Exercise 303 *Do Exercise for $m = 3$, $k = 12$, and $\gamma = 12$.*

Exercise 304 *Using the mks units (meters-kilograms-seconds), suppose you have a spring with spring constant $4\,\mathrm{N/m}$. You want to use it to weigh items. Assume no friction. You place the mass on the spring and put it in motion.*

a) *You count and find that the frequency is 0.8 Hz (cycles per second). What is the mass? (Be careful with the units here, the frequency is given in cycles per second, not radians per second.)*

b) *Find a formula for the mass m given the frequency ω in Hz.*

Exercise 305 *A mass of 2 kilograms is on a spring with spring constant k newtons per meter with no damping. Suppose the system is at rest and at time $t = 0$ the mass is kicked and starts traveling at 2 meters per second. How large does k have to be to so that the mass does not go further than 3 meters from the rest position?*

Exercise 306 *Suppose we add possible friction to Exercise . Further, suppose you do not know the spring constant, but you have two reference weights 1 kg and 2 kg to calibrate your setup. You put each in motion on your spring and measure the quasi-frequency. For the 1 kg weight you measured 1.1 Hz, for the 2 kg weight you measured 0.8 Hz.*

a) *Find k (spring constant) and γ (damping constant).*

b) *Find a formula for the mass in terms of the frequency in Hz.* Note that there may be more than one possible mass for a given frequency.

c) *For an unknown object you measured 0.2 Hz, what is the mass of the object? Suppose that you know that the mass of the unknown object is more than a kilogram.*

Exercise 307 *Suppose you wish to measure the friction a mass of 0.1 kg experiences as it slides along a floor (you wish to find γ). You have a spring with spring constant $k = 5\,\mathrm{N/m}$. You take the spring, you attach it to the mass and fix it to a wall. Then you pull on the spring and let the mass go. You find that the mass oscillates with quasi-frequency 1 Hz. What is the friction?*

Author(s): Matthew Charnley and Jason Nowell

Exercise 308 *A 5000 kg railcar hits a bumper (a spring) at 1 m/s, and the spring compresses by 0.1 m. Assume no damping.*

a) *Find k.*

b) *How far does the spring compress when a 10000 kg railcar hits the spring at the same speed?*

c) *If the spring would break if it compresses further than 0.3 m, what is the maximum mass of a railcar that can hit it at 1 m/s?*

d) *What is the maximum mass of a railcar that can hit the spring without breaking at 2 m/s?*

Exercise 309 *When attached to a spring, a 2 kg mass stretches the spring by 0.49 m.*

a) *What is the spring constant of this spring? Use 9.8 m/s² as the gravity constant.*

b) *This mass is allowed to come to rest, lifted up by 0.4 m and then released. If there is no damping, set up and solve an initial value problem for the position of the mass as a function of time.*

c) *For a next experiment, you attach a dampener of coefficient 16 Ns/m to the system, and give the same initial condition. Set up and solve an initial value problem for the position of the mass. What type of "dampening" would be used to characterize this situation?*

Exercise 310 *A mass of m kg is on a spring with $k = 3\,\mathrm{N/m}$ and $c = 2\,\mathrm{Ns/m}$. Find the mass m_0 for which there is critical damping. If $m < m_0$, does the system oscillate or not, that is, is it underdamped or overdamped?*

Exercise 311 *Suppose we have an RLC circuit with a resistor of 100 milliohms (0.1 ohms), inductor of inductance of 50 millihenries (0.05 henries), and a capacitor of 5 farads, with constant voltage.*

a) *Set up the ODE equation for the current I.*

b) *Find the general solution.*

c) *Solve for $I(0) = 10$ and $I'(0) = 0$.*

Exercise 312 *For RLC circuits, we can use either charge or current to set up the equation. Let's see how the two of those compare.*

a) *Assume that we have an RLC circuit with a 30 millihenry inductor, a 120 milliohm resistor, and a capacitor with capacitance $20/3$ F. Set up a differential equation for the charge on the capacitor as a function of time.*

b) *Use the same circuit to set up a differential equation for the current through the circuit as a function of time. How do these equations relate?*

c) *Find the general solution to each of these equations.*

d) *Solve the initial value problem for the charge with $Q(0) = 1/2C$ and $Q'(0) = 0$.*

e) *Using the fact that $I = Q'$, determine the appropriate initial conditions needed for I in order to solve for the current in this same setup (with those initial values for charge).*

f) *Now, we'll do the same in the other direction. Solve the initial value problem for current with $I(0) = 2A$ and $I'(0) = 1\,\mathrm{A/s}$, and see what the initial conditions would be for $Q(t)$ for this setup.*

Exercise 313 *Assume that the system $my'' + \gamma y' + ky = 0$ is either critically or overdamped. Prove that the solution can pass through zero* at most once, *regardless of initial conditions.* Hint: *Try to find all values of t for which $y(t) = 0$, given the form of the solution.*

Nonhomogeneous equations

We discuss Nonhomogeneous systems

Solving nonhomogeneous equations

We have solved linear constant coefficient homogeneous equations. What about nonhomogeneous linear ODEs? For example, the equations for forced mechanical vibrations, where we add a "forcing" term, which is a function on the right-hand side of the equation. That is, suppose we have an equation such as

$$y'' + 5y' + 6y = 2x + 1. \tag{22}$$

We will write $L[y] = 2x + 1$, where $L[y]$ represents the entire left-hand side of $y'' + 5y' + 6y$, when the exact form of the operator is not important. We solve (22) in the following manner. First, we find the general solution y_c to the *associated homogeneous equation*

$$y'' + 5y' + 6y = 0. \tag{23}$$

We call y_c the *complementary solution*. Next, we find a single *particular solution* y_p to (22) in some way (that is the point of this section). Then

$$y = y_c + y_p$$

is the general solution to (22). We have $L[y_c] = 0$ and $L[y_p] = 2x + 1$. As L is a *linear operator* we verify that y is a solution, $L[y] = L[y_c + y_p] = L[y_c] + L[y_p] = 0 + (2x + 1)$. Let us see why we obtain the *general* solution.

Let y_p and $\tilde{y}_p$ be two different particular solutions to (22). Write the difference as $w = y_p - \tilde{y}_p$. Then plug w into the left-hand side of the equation to get

$$w'' + 5w' + 6w = (y_p'' + 5y_p' + 6y_p) - (\tilde{y}_p'' + 5\tilde{y}_p' + 6\tilde{y}_p) = (2x + 1) - (2x + 1) = 0.$$

Using the operator notation the calculation becomes simpler. As L is a linear operator we write

$$L[w] = L[y_p - \tilde{y}_p] = L[y_p] - L[\tilde{y}_p] = (2x + 1) - (2x + 1) = 0.$$

So $w = y_p - \tilde{y}_p$ is a solution to (23), that is $Lw = 0$. However, we know what all solutions to $Lw = 0$ look like, as this is a homogeneous equation that we have solved previously. Therefore, any two solutions of (22) differ by a solution to the homogeneous equation (23). The solution $y = y_c + y_p$ includes *all* solutions to (22), since y_c is the general solution to the associated homogeneous equation.

Theorem 10. *Let $L[y] = f(x)$ be a linear ODE (not necessarily constant coefficient). Let y_c be the complementary solution (the general solution to the associated homogeneous equation $L[y] = 0$) and let y_p be any particular solution to $L[y] = f(x)$. Then the general solution to $L[y] = f(x)$ is*

$$y = y_c + y_p.$$

The moral of the story is that we can find the particular solution in any old way. If we find a different particular solution (by a different method, or simply by guessing), then we still get the same general solution. The formula may look different, and the constants we have to choose to satisfy the initial conditions may be different, but it is the same solution.

Undetermined coefficients

The trick is to somehow, in a smart way, guess one particular solution to (22). Note that $2x + 1$ is a polynomial, and the left-hand side of the equation (with all of the derivatives) will still be a polynomial if we let y be a polynomial of the same degree. Let us try

$$y_p = Ax + B.$$

Learning outcomes: Find the corresponding homogeneous equation for a non-homogeneous equation Use the method of undetermined coefficients to solve non-homogeneous equations Use variation of parameters to solve non-homogeneous equations Solve for the necessary coefficients to solve initial value problems for non-homogeneous equations.

Author(s): Matthew Charnley and Jason Nowell

We plug y_p into the left hand side to obtain

$$\begin{aligned} y_p'' + 5y_p' + 6y_p &= (Ax+B)'' + 5(Ax+B)' + 6(Ax+B) \\ &= 0 + 5A + 6Ax + 6B = 6Ax + (5A+6B). \end{aligned}$$

So $6Ax + (5A+6B) = 2x+1$. If we match up the coefficients of x in this equation, we get that $6A = 2$ or $A = 1/3$. In order for the constant terms to match, we need that $5A + 6B = 1$. Since we know the value of A, this tells us that $B = -1/9$. That means $y_p = \frac{1}{3}x - \frac{1}{9} = \frac{3x-1}{9}$. Solving the complementary problem (exercise!) we get

$$y_c = C_1 e^{-2x} + C_2 e^{-3x}.$$

Hence the general solution to (22) is

$$y = C_1 e^{-2x} + C_2 e^{-3x} + \frac{3x-1}{9}.$$

Now suppose we are further given some initial conditions. For example, $y(0) = 0$ and $y'(0) = 1/3$. First find $y' = -2C_1 e^{-2x} - 3C_2 e^{-3x} + 1/3$. Then

$$0 = y(0) = C_1 + C_2 - \frac{1}{9}, \qquad \frac{1}{3} = y'(0) = -2C_1 - 3C_2 + \frac{1}{3}.$$

We solve to get $C_1 = 1/3$ and $C_2 = -2/9$. The particular solution we want is

$$y(x) = \frac{1}{3}e^{-2x} - \frac{2}{9}e^{-3x} + \frac{3x-1}{9} = \frac{3e^{-2x} - 2e^{-3x} + 3x - 1}{9}.$$

Exercise 314 *Check that y really solves the equation (22) and the given initial conditions.*

Note: A common mistake is to solve for constants using the initial conditions with y_c and only add the particular solution y_p after that. That will *not* work. You need to first compute $y = y_c + y_p$ and *only then* solve for the constants using the initial conditions.

A right-hand side consisting of exponentials, sines, and cosines can be handled similarly.

Example 57. *One example of this is*

$$y'' + 2y' + 2y = \cos(2x).$$

Solution: Let us find some y_p. We start by guessing that the solution includes some multiple of $\cos(2x)$. We try

$$y_p = A\cos(2x).$$

Plugging this into the differential equation gives

$$\underbrace{-4A\cos(2x)}_{y_p''} + 2\underbrace{\big(-2A\sin(2x)\big)}_{y_p'} + 2\underbrace{\big(A\cos(2x)\big)}_{y_p} = \cos(2x).$$

Simplifying this expression gives

$$-2A\cos(2x) - 4A\sin(2x) = \cos(2x)$$

and we have a problem. Since there is no sine term on the right-hand side, we are forced to pick $A = 0$, which means our non-homogeneous solution is zero, and that's not good. What happened here? In the previous example, when we differentiated a polynomial (as part of the y_p guess) the function stayed a polynomial, and so we did not add any new types of terms. In this case, however, when we differentiate the cosine term in our guess, it becomes a sine, which we did *not* have in our initial guess.

Thus, we will also want to add a multiple of $\sin(2x)$ to our guess since derivatives of cosine are sines. We try

$$y_p = A\cos(2x) + B\sin(2x).$$

We plug y_p into the equation and we get

$$\begin{aligned} &\underbrace{-4A\cos(2x) - 4B\sin(2x)}_{y_p''} + 2\underbrace{\big(-2A\sin(2x) + 2B\cos(2x)\big)}_{y_p'} \\ &\qquad + 2\underbrace{\big(A\cos(2x) + B\sin(2x)\big)}_{y_p} = \cos(2x), \end{aligned}$$

or

$$(-4A + 4B + 2A)\cos(2x) + (-4B - 4A + 2B)\sin(2x) = \cos(2x).$$

The left-hand side must equal to right-hand side. Namely, $-4A + 4B + 2A = 1$ and $-4B - 4A + 2B = 0$. So $-2A + 4B = 1$ and $2A + B = 0$. We can solve this system of equations to get that $A = -1/10$ and $B = 1/5$. So

$$y_p = A\cos(2x) + B\sin(2x) = \frac{-\cos(2x) + 2\sin(2x)}{10}.$$

Similarly, if the right-hand side contains exponentials we try exponentials. If

$$L[y] = e^{3x},$$

we try $y = Ae^{3x}$ as our guess and try to solve for A.

When the right-hand side is a multiple of sines, cosines, exponentials, and polynomials, we can use the product rule for differentiation to come up with a guess. We need to guess a form for y_p such that $L[y_p]$ is of the same form, and has all the terms needed to for the right-hand side. For example,

$$L[y] = (1 + 3x^2)\, e^{-x} \cos(\pi x).$$

For this equation, we guess

$$y_p = (A + Bx + Cx^2)\, e^{-x} \cos(\pi x) + (D + Ex + Fx^2)\, e^{-x} \sin(\pi x).$$

We plug in and then hopefully get equations that we can solve for A, B, C, D, E, and F. As you can see this can make for a very long and tedious calculation very quickly. C'est la vie!

There is one hiccup in all this. It could be that our guess actually solves the associated homogeneous equation. That is, suppose we have

$$y'' - 9y = e^{3x}.$$

We would love to guess $y = Ae^{3x}$, but if we plug this into the left-hand side of the equation we get

$$y'' - 9y = 9Ae^{3x} - 9Ae^{3x} = 0 \neq e^{3x}.$$

There is no way we can choose A to make the left-hand side be e^{3x}. The trick in this case is to multiply our guess by x to get rid of duplication with the complementary solution. That is first we compute y_c (solution to $L[y] = 0$)

$$y_c = C_1 e^{-3x} + C_2 e^{3x},$$

and we note that the e^{3x} term is a duplicate with our desired guess. We modify our guess to $y = Axe^{3x}$ so that there is no duplication anymore. Let us try: $y' = Ae^{3x} + 3Axe^{3x}$ and $y'' = 6Ae^{3x} + 9Axe^{3x}$, so

$$y'' - 9y = 6Ae^{3x} + 9Axe^{3x} - 9Axe^{3x} = 6Ae^{3x}.$$

Thus $6Ae^{3x}$ is supposed to equal e^{3x}. Hence, $6A = 1$ and so $A = 1/6$. We can now write the general solution as

$$y = y_c + y_p = C_1 e^{-3x} + C_2 e^{3x} + \frac{1}{6}\, xe^{3x}.$$

Notice that the term of the form xe^{3x} does not show up on the left-hand side after differentiating the equation, and the only term that survives is the e^{3x} term that showed up from the derivatives. This works out because e^{3x} solves the homogeneous problem. With that though, make sure to remember to include the xe^{3x} when you write out the general solution at the end of the problem, because it does appear there.

It is possible that multiplying by x does not get rid of all duplication. For example,

$$y'' - 6y' + 9y = e^{3x}.$$

The complementary solution is $y_c = C_1 e^{3x} + C_2 xe^{3x}$. Guessing $y = Axe^{3x}$ would not get us anywhere. In this case we want to guess $y_p = Ax^2 e^{3x}$. Basically, we want to multiply our guess by x until all duplication is gone. *But no more!* Multiplying too many times will not work (in that case, the derivatives won't actually get down to the plain e^{3x} term that you need in order to solve the problem).

Finally, what if the right-hand side has several terms, such as

$$L[y] = e^{2x} + \cos x.$$

In this case we find u that solves $L[u] = e^{2x}$ and v that solves $L[v] = \cos x$ (that is, do each term separately). Then note that if $y = u + v$, then $L[y] = e^{2x} + \cos x$. This is because L is linear; we have $L[y] = L[u + v] = L[u] + L[v] = e^{2x} + \cos x$.

To summarize all of this, we can make a table of the different guesses we should make given the form of the right hand side.

Right hand side	**Guess**
$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$	$Ax^n + Bx^{n-1} + \cdots + Nx + P$
e^{ax}	Ae^{ax}
$\cos ax$	$A\cos ax + B\sin ax$
$\sin ax$	$A\cos ax + B\sin ax$

- If there is a product of above terms, guess the product of the guesses. So, for a right hand side of xe^{ax}, the guess should be $(Ax + B)e^{ax}$, and for a right hand side of $x\cos ax$, the guess should be $(Ax + B)\cos ax + (Cx + D)\sin ax$.
- If any part solves the homogeneous problem, multiply that entire component by x until nothing does.

Example 58. *Find the solution to the initial value problem*

$$y'' - 3y' - 4y = 2e^{-x} + 4\sin(x) \qquad y(0) = -2,\ y'(0) = 1$$

Solution: To start this problem, we look for the solution to the homogeneous problem. The characteristic equation for the left hand side is $r^2 - 3r - 4$, which factors as $(r - 4)(r + 1)$. Therefore the general solution to the homogeneous problem (or the complementary solution) is

$$y_c(x) = C_1 e^{4x} + C_2 e^{-x}.$$

Next, we want to use undetermined coefficients to solve the non-homogeneous problem. Note that we have to wait until after this part to meet the initial conditions. Since our right-hand side is $2e^{-x} + 4\sin(x)$, we need to guess two components for the two different terms in this function. For the first term, we would want to guess Ae^{-x}, but this function solves the homogeneous problem. Therefore, we need to multiply by x to use Axe^{-x} as our guess. For the sine term, we need to guess both sine and cosine, so we add $B\sin(x) + C\cos(x)$ to our guess. Therefore, our total guess for the non-homogeneous solution is

$$y_p(x) = Axe^{-x} + B\sin(x) + C\cos(x).$$

We take two derivatives of this function and then plug it into the differential equation

$$\begin{aligned} y_p(x) &= Axe^{-x} + B\sin(x) + C\cos(x) \\ y_p'(x) &= Ae^{-x} - Axe^{-x} + B\cos(x) - C\sin(x) \\ y_p''(x) &= Axe^{-x} - 2Ae^{-x} - B\sin(x) - C\cos(x) \end{aligned}$$

so that

$$\begin{aligned} y_p'' - 3y_p' - 4y_p = (&Axe^{-x} - 2Ae^{-x} - B\sin(x) - C\cos(x)) \\ &- 3(Ae^{-x} - Axe^{-x} + B\cos(x) - C\sin(x)) \\ &- 4(Axe^{-x} + B\sin(x) + C\cos(x)) \end{aligned}$$

which can be simplified to

$$y_p'' - 3y_p' - 4y_p = -5Ae^{-x} + (3C - 5B)\sin(x) + (-3B - 5C)\cos(x).$$

Since we want this to equal $2e^{-x} + 4\sin(x)$, this means that we need $-5A = 2$, so $A = -2/5$, as well as $3C - 5B = 4$ and $-3B - 5C = 0$. The second of these implies that $3B = -5C$, or $B = -5/3C$, so that the first equation gives $3C - 5(-5/3C) = 4$. This implies that $(3 + 25/3)C = 4$ so that

$$C = \frac{4}{(3 + \frac{25}{3})} = \frac{4}{\frac{34}{3}} = \frac{6}{17}.$$

We can then find B as

$$B = -\frac{5}{3}C = -\frac{5}{3} \cdot \frac{6}{17} = -\frac{10}{17}.$$

Therefore, the general solution to this non-homogeneous problem is

$$y(x) = C_1 e^{4x} + C_2 e^{-x} - \frac{2}{5} x e^{-x} - \frac{10}{17}\sin(x) + \frac{6}{17}\cos(x).$$

Now we can look to meet the initial conditions. We want to differentiate this expression to get

$$y'(x) = 4C_1 e^{4x} - C_2 e^{-x} - \frac{2}{5}e^{-x} + \frac{2}{5}xe^{-x} - \frac{10}{17}\cos(x) - \frac{6}{17}\sin(x)$$

and then plug zero into both y and y' to get that

$$\begin{aligned} y(0) &= C_1 + C_2 + \frac{6}{17} = -2 \\ y'(0) &= 4C_1 - C_2 - \frac{2}{5} - \frac{10}{17} = 1 \end{aligned}$$

which gives rise to the system

$$C_1 + C_2 = -\frac{40}{17} \qquad 4C_1 - C_2 = \frac{169}{85}.$$

Adding the equations together gives $5C_1 = -\frac{31}{85}$ so that $C_1 = -\frac{31}{425}$ and then $C_2 = -\frac{969}{425} = -\frac{57}{25}$. Therefore the solution to the initial value problem is

$$y(x) = -\frac{31}{425}e^{4x} - \frac{57}{25}e^{-x} - \frac{2}{5}xe^{-x} - \frac{10}{17}\sin(x) + \frac{6}{17}\cos(x).$$

Exercise 315 *Verify that this $y(x)$ solves the initial value problem!*

Variation of parameters

The method of undetermined coefficients works for many basic problems that crop up. But it does not work all the time. It only works when the right-hand side of the equation $L[y] = f(x)$ has finitely many linearly independent derivatives, so that we can write a guess that consists of them all. Some equations are a bit tougher. Consider

$$y'' + y = \tan x.$$

Each new derivative of $\tan x$ looks completely different and cannot be written as a linear combination of the previous derivatives. If we start differentiating $\tan x$, we get:

$$\sec^2 x, \quad 2\sec^2 x \tan x, \quad 4\sec^2 x \tan^2 x + 2\sec^4 x,$$
$$8\sec^2 x \tan^3 x + 16\sec^4 x \tan x, \quad 16\sec^2 x \tan^4 x + 88\sec^4 x \tan^2 x + 16\sec^6 x, \quad \ldots$$

This equation calls for a different method. We present the method of *variation of parameters*, which handles any equation of the form $L[y] = f(x)$, provided we can solve certain integrals. For simplicity, we restrict ourselves to second order constant coefficient equations, but the method works for higher order equations just as well (the computations become more tedious). The method also works for equations with nonconstant coefficients, provided we can solve the associated homogeneous equation.

Perhaps it is best to explain this method by example. Let us try to solve the equation

$$L[y] = y'' + y = \tan x.$$

First we find the complementary solution (solution to $L[y_c] = 0$). We get $y_c = C_1 y_1 + C_2 y_2$, where $y_1 = \cos x$ and $y_2 = \sin x$. To find a particular solution to the nonhomogeneous equation we try

$$y_p = y = u_1 y_1 + u_2 y_2,$$

where u_1 and u_2 are *functions* and not constants. We are trying to satisfy $L[y] = \tan x$. That gives us one condition on the functions u_1 and u_2. Compute (note the product rule!)

$$y' = (u_1' y_1 + u_2' y_2) + (u_1 y_1' + u_2 y_2').$$

We can still impose one more condition at our discretion to simplify computations (we have two unknown functions, so we should be allowed two conditions). We require that $(u_1' y_1 + u_2' y_2) = 0$. This makes computing the second derivative easier.

$$\begin{aligned} y' &= u_1 y_1' + u_2 y_2', \\ y'' &= (u_1' y_1' + u_2' y_2') + (u_1 y_1'' + u_2 y_2''). \end{aligned}$$

Since y_1 and y_2 are solutions to $y'' + y = 0$, we find $y_1'' = -y_1$ and $y_2'' = -y_2$. (If the equation was a more general $y'' + p(x)y' + q(x)y = 0$, we would have $y_i'' = -p(x)y_i' - q(x)y_i$.) So

$$y'' = (u_1' y_1' + u_2' y_2') - (u_1 y_1 + u_2 y_2).$$

We have $(u_1 y_1 + u_2 y_2) = y$ and so

$$y'' = (u_1' y_1' + u_2' y_2') - y,$$

and hence

$$y'' + y = L[y] = u_1' y_1' + u_2' y_2'.$$

For y to satisfy $L[y] = f(x)$ we must have $f(x) = u_1' y_1' + u_2' y_2'$.

What we need to solve are the two equations (conditions) we imposed on u_1 and u_2:

$$\begin{aligned} u_1' y_1 + u_2' y_2 &= 0, \\ u_1' y_1' + u_2' y_2' &= f(x). \end{aligned}$$

We solve for u_1' and u_2' in terms of $f(x)$, y_1 and y_2. We always get these formulas for any $L[y] = f(x)$, where $L[y] = y'' + p(x)y' + q(x)y$. There is a general formula for the solution we could just plug into, but instead of memorizing that, it is better, and easier, to just repeat what we do below. In our case the two equations are

$$\begin{aligned} u_1' \cos(x) + u_2' \sin(x) &= 0, \\ -u_1' \sin(x) + u_2' \cos(x) &= \tan(x). \end{aligned}$$

Hence

$$\begin{aligned} u_1' \cos(x)\sin(x) + u_2' \sin^2(x) &= 0, \\ -u_1' \sin(x)\cos(x) + u_2' \cos^2(x) &= \tan(x)\cos(x) = \sin(x). \end{aligned}$$

And thus

$$\begin{aligned} & u_2' \bigl(\sin^2(x) + \cos^2(x)\bigr) = \sin(x), \\ & u_2' = \sin(x), \\ & u_1' = \frac{-\sin^2(x)}{\cos(x)} = -\tan(x)\sin(x). \end{aligned}$$

We integrate u_1' and u_2' to get u_1 and u_2.

$$\begin{aligned} u_1 &= \int u_1'\,dx = \int -\tan(x)\sin(x)\,dx = \frac{1}{2}\ln\left|\frac{\sin(x)-1}{\sin(x)+1}\right| + \sin(x), \\ u_2 &= \int u_2'\,dx = \int \sin(x)\,dx = -\cos(x). \end{aligned}$$

So our particular solution is

$$\begin{aligned} y_p = u_1 y_1 + u_2 y_2 = \frac{1}{2}\cos(x)\ln\left|\frac{\sin(x)-1}{\sin(x)+1}\right| + \cos(x)\sin(x) - \cos(x)\sin(x) = \\ = \frac{1}{2}\cos(x)\ln\left|\frac{\sin(x)-1}{\sin(x)+1}\right|. \end{aligned}$$

The general solution to $y'' + y = \tan x$ is, therefore,

$$y = C_1\cos(x) + C_2\sin(x) + \frac{1}{2}\cos(x)\ln\left|\frac{\sin(x)-1}{\sin(x)+1}\right|.$$

In more generality, we can take the system of equations

$$\begin{aligned} u_1'y_1 + u_2'y_2 &= 0, \\ u_1'y_1' + u_2'y_2' &= f(x). \end{aligned}$$

and solve out for u_1' and u_2' using elimination. If we do that, we get that

$$u_1' = -\frac{y_2(x)f(x)}{y_1(x)y_2'(x) - y_1'(x)y_2(x)} \qquad u_2' = \frac{y_1(x)f(x)}{y_1(x)y_2'(x) - y_1'(x)y_2(x)}.$$

We know that solving the equations this way will work out because we start with the assumption that y_1 and y_2 are linearly independent solutions, and the denominator of both of these fractions is exactly what we know is not zero from this assumption. Therefore, both of these functions can be written this way, we can integrate both of them, and set up our particular solution of the form $y_p(x) = u_1y_1 + u_2y_2$ to get

$$y_p(x) = -y_1(x)\int_{x_0}^{x}\frac{y_2(r)f(r)}{y_1(r)y_2'(r) - y_1'(r)y_2(r)}\,dr + y_2(x)\int_{x_0}^{x}\frac{y_1(r)f(r)}{y_1(r)y_2'(r) - y_1'(r)y_2(r)}\,dr \tag{24}$$

where x_0 is any conveniently chosen value (usually zero). Notice the use of r as a dummy variable here to separate the functions being integrated from the actual variable that shows up in the solution. This formula will always work for finding a particular solution to a non-homogeneous equation given that we know the solution to the homogeneous equation, but we may not be able to work out the integrals explicitly. This is the downside of this method, it may always work, but can be very tedious and may not result in nice, closed-form expressions like we might get from other methods.

Example 59. *Find the general solution to the differential equation*

$$y'' + 4y' + 3y = e^{3x} + 2$$

using both undetermined coefficients and variation of parameters.

Solution: For both methods of solving non-homogeneous equations, we need the solution to the homogeneous problem. For this equation, the characteristic polynomial is $r^2 + 4r + 3$, which factors as $(r+1)(r+3)$, so the general solution to the homogeneous problem is

$$y_c(x) = C_1e^{-x} + C_2e^{-3x}.$$

To use undetermined coefficients, we need to get the appropriate guess for the right-hand side, which in this case is $y_p(x) = Ae^{3x} + B$. Plugging this in to the differential equation gives

$$9Ae^{3x} + 4(3Ae^{3x}) + 3(Ae^{3x} + B) = e^{3x} + 2$$

which simplifies to

$$24Ae^{3x} + 3B = e^{3x} + 2$$

so that $A = 1/24$ and $B = 2/3$. Thus, the general solution to the non-homogeneous equation is

$$y(x) = C_1e^{-x} + C_2e^{-3x} + \frac{1}{24}e^{3x} + \frac{2}{3}.$$

In order to use variation of parameters, we let $y_1(x) = e^{-x}$ and $y_2(x) = e^{-3x}$ be the two linearly independent solutions that we found to the homogeneous problem. Our right-hand side function is $f(x) = e^{3x} + 2$ and we can compute the expression

$$y_1(x)y_2'(x) - y_1'(x)y_2(x) = e^{-x}(-3e^{-3x}) - (-e^{-x})e^{-3x} = -2e^{-4x}.$$

Therefore, we can use the formulas from the method of variation of parameters to compute that

$$\begin{aligned} u_1' &= -\frac{y_2(x)f(x)}{y_1(x)y_2'(x) - y_1'(x)y_2(x)} = -\frac{e^{-3x}(e^{3x}+2)}{-2e^{-4x}} = \frac{1}{2}e^{4x} + e^{x} \\ u_2' &= \frac{y_1(x)f(x)}{y_1(x)y_2'(x) - y_1'(x)y_2(x)} = \frac{e^{-x}(e^{3x}+2)}{-2e^{-4x}} = -\frac{1}{2}e^{6x} - e^{3x}. \end{aligned}$$

Then we can compute

$$u_1 = \frac{1}{8}e^{4x} + e^x + C_1 \qquad u_2 = -\frac{1}{12}e^{6x} - \frac{1}{3}e^{3x} + C_2.$$

Then, we can write out the full general solution as $y(x) = u_1(x)y_1(x) + u_2(x)y_2(x)$ or

$$\begin{aligned} y(x) &= e^{-x}\left(\frac{1}{8}e^{4x} + e^x + C_1\right) + e^{-3x}\left(-\frac{1}{12}e^{6x} - \frac{1}{3}e^{3x} + C_2\right) \\ &= \frac{1}{8}e^{3x} + 1 + C_1e^{-x} - \frac{1}{12}e^{3x} - \frac{1}{3} + C_2e^{-3x} \end{aligned}$$

which, after combining the terms, is the same as the solution that we obtained via undetermined coefficients. $\lrcorner$

higherOrder/practice/nonHomogenEqns-practice1.tex

Practice for Non-Homogeneous Equations

Why?

Exercise 316 *Find a particular solution of* $y'' - y' - 6y = e^{2x}$.

Exercise 317 *Find a particular solution of* $y'' - 4y' + 4y = e^{2x}$.

Exercise 318 *Find a particular solution to* $y'' - y' + y = 2\sin(3x)$

Exercise 319 *Solve the initial value problem* $y'' + 9y = \cos(3x) + \sin(3x)$ *for* $y(0) = 2$, $y'(0) = 1$.

Exercise 320 *Set up the form of the particular solution but do not solve for the coefficients for* $y^{(4)} - 2y''' + y'' = e^x$.

Exercise 321 *Set up the form of the particular solution but do not solve for the coefficients for* $y^{(4)} - 2y''' + y'' = e^x + x + \sin x$.

Exercise 322 *Solve* $y'' + 2y' + y = x^2$, $y(0) = 1$, $y'(0) = 2$.

Exercise 323 *Use the method of undetermined coefficients to solve the DE* $y'' + 4y' = 2t + 30$.

Exercise 324

a) *Using variation of parameters find a particular solution of* $y'' - 2y' + y = e^x$.

b) *Find a particular solution using undetermined coefficients.*

c) *Are the two solutions you found the same? See also* Exercise .

Exercise 325

a) *Find a particular solution to* $y'' + 2y = e^x + x^3$.

b) *Find the general solution.*

Exercise 326 *Find the general solution to* $y'' - 3y' - 4y = e^{2t} + 1$.

Author(s): Matthew Charnley and Jason Nowell

Exercise 327 *Find the general solution to* $y'' - 2y' + 5y = \sin(3t) + 2\cos(3t)$.

Exercise 328 *Find the general solution to* $y'' - 4y' - 21y = e^{-3t} + e^{4t}$.

Exercise 329 *Find the general solution to* $y'' - 2y' + y = e^t - t$.

Exercise 330 *Find the general solution to* $y'' + 4y = \sec(2t)$ *using variation of parameters.*

Exercise 331 *Find the solution of the initial value problem* $y'' - 2y' - 15y = e^{5t} + 3$, $y(0) = 2$, $y'(0) = -1$.

Exercise 332 *Find the solution of the initial value problem* $y'' + 4y' + 5y = \cos(3t) + t$, $y(0) = 0$, $y'(0) = 2$.

Exercise 333 *The following differential equations are all related. Find the general solution to each of them and compare and contrast the different solutions and the methods used to approach them.*

a) $y'' - 2y' - 15y = e^t + 5e^{-4t}$

b) $y'' - 2y' - 15y = 2e^{2t} + 3e^{-t}$

c) $y'' - 2y' - 15y = 3\cos(2t)$

d) $y'' - 2y' - 15y = 2e^{5t} - \sin(t)$

Exercise 334 *The following differential equations are all related. Find the general solution to each of them and compare and contrast the different solutions and the methods used to approach them.*

a) $y'' + 4y' + 3y = e^{2t} + 3e^{4t}$

b) $y'' - 2y' + 5y = e^{2t} + 3e^{4t}$

c) $y'' + 3y' - 10y = e^{2t} + 3e^{4t}$

d) $y'' - 8y' + 16y = e^{2t} + 3e^{4t}$

Exercise 335 *Find a particular solution of* $y'' - 2y' + y = \sin(x^2)$. *It is OK to leave the answer as a definite integral.*

Exercise 336 *Use variation of parameters to find a particular solution of* $y'' - y = \dfrac{1}{e^x + e^{-x}}$.

Exercise 337 *Recall that a homogeneous* Euler equation *is one of the form* $t^2y'' + aty' + by = 0$ *and is solved by using the guess* $y(t) = t^r$ *and solving for the potential values of* r.

a) *Solve* $t^2y'' - 2ty' - 10y = 0$.

b) *Let* y_1 *and* y_2 *be a fundamental set for the above equation. Use the variation of parameters equations* $u_1 = -\int \frac{y_2\, g(t)}{y_1y_2' - y_2y_1'}\, dt$, $y_2 = \int \frac{y_1\, g(t)}{y_1y_2' - y_2y_1'}\, dt$ *to solve the non-homogeneous equation* $y'' - \frac{2}{t} - \frac{10}{t^2} = t^3$.
(Do not attempt method of undetermined coefficients instead; it won't work.)

Exercise 338 *For an arbitrary constant* c *find the general solution to* $y'' - 2y = \sin(x + c)$.

Exercise 339 *For an arbitrary constant* c *find a particular solution to* $y'' - y = e^{cx}$. *Hint: Make sure to handle every possible real* c.

Exercise 340

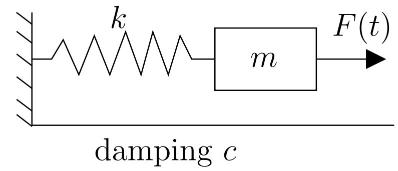
a) *Using variation of parameters find a particular solution of* $y'' - y = e^x$.

b) *Find a particular solution using undetermined coefficients.*

c) *Are the two solutions you found the same? What is going on?*

Forced oscillations and resonance

We discuss Forced oscillations and resonance



Let us return back to the example of a mass on a spring. We examine the case of forced oscillations, which we did not yet handle. That is, we consider the equation

$$mx'' + \gamma x' + kx = F(t),$$

for some nonzero $F(t)$. The setup is again: m is mass, γ is friction, k is the spring constant, and $F(t)$ is an external force acting on the mass.

We are interested in periodic forcing, such as noncentered rotating parts, or perhaps loud sounds, or other sources of periodic force.

Undamped forced motion and resonance

First let us consider undamped ($\gamma = 0$) motion. We have the equation

$$mx'' + kx = F_0 \cos(\omega t).$$

This equation has the complementary solution (solution to the associated homogeneous equation)

$$x_c = C_1 \cos(\omega_0 t) + C_2 \sin(\omega_0 t),$$

where $\omega_0 = \sqrt{k/m}$ is the *natural frequency* (angular). It is the frequency at which the system "wants to oscillate" without external interference.

Suppose that $\omega_0 \neq \omega$. We try the solution $x_p = A\cos(\omega t)$ and solve for A. We do not need a sine in our trial solution as after plugging in we only have cosines. If you include a sine, it is fine; you will find that its coefficient is zero (I could not find a second rhyme).

We solve using the method of undetermined coefficients. We find that

$$x_p = \frac{F_0}{m(\omega_0^2 - \omega^2)} \cos(\omega t).$$

We leave it as an exercise to do the algebra required.

The general solution is

$$x = C_1 \cos(\omega_0 t) + C_2 \sin(\omega_0 t) + \frac{F_0}{m(\omega_0^2 - \omega^2)} \cos(\omega t).$$

Written another way

$$x = C \cos(\omega_0 t - \gamma) + \frac{F_0}{m(\omega_0^2 - \omega^2)} \cos(\omega t).$$

The solution is a superposition of two cosine waves at different frequencies.

Example 60. *Take*

$$0.5x'' + 8x = 10\cos(\pi t), \qquad x(0) = 0, \qquad x'(0) = 0.$$

Solution: Let us compute. First we read off the parameters: $\omega = \pi$, $\omega_0 = \sqrt{8/0.5} = 4$, $F_0 = 10$, $m = 0.5$. The general solution is

$$x = C_1 \cos(4t) + C_2 \sin(4t) + \frac{20}{16 - \pi^2} \cos(\pi t).$$

Learning outcomes: Write differential equations to model forced oscillators (like masses on springs) Identify when beats, pure resonance, and practical resonance can occur Use proper terminology around transient and steady periodic solutions when discussing these problems.

Author(s): Matthew Charnley and Jason Nowell

Solve for C_1 and C_2 using the initial conditions: $C_1 = \dfrac{-20}{16-\pi^2}$ and $C_2 = 0$. Hence

$$x = \frac{20}{16-\pi^2}\big(\cos(\pi t) - \cos(4t)\big).$$

Notice the "beating" behavior in **Normally a reference to a previous figure goes here.**. First use the trigonometric identity

$$2\sin\left(\frac{A-B}{2}\right)\sin\left(\frac{A+B}{2}\right) = \cos B - \cos A$$

to get

$$x = \frac{20}{16-\pi^2}\left(2\sin\left(\frac{4-\pi}{2}t\right)\sin\left(\frac{4+\pi}{2}t\right)\right).$$

The function x is a high frequency wave modulated by a low frequency wave.

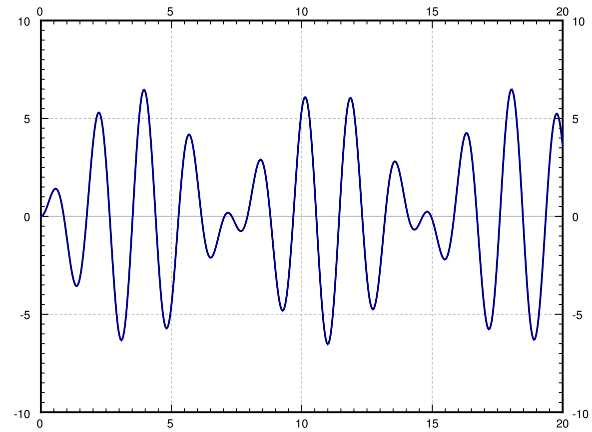


Figure 45: Graph of $\dfrac{20}{16-\pi^2}\big(\cos(\pi t) - \cos(4t)\big)$.

The beating behavior can be experienced even more readily by considering a higher frequency and viewing the resulting function as a sound wave. A sound wave of frequency 440 Hz produces and A4 sound, which is the A above middle C on a piano. This means that the function

$$x_p(t) = \sin(2\pi \cdot 440t)$$

will produce a sound wave equivalent to this A4 sound. In MATLAB, this can be done with the code

Code

```
omega0 = 440*2*pi;
tVals = linspace(0, 5, 5*8192);

testSound = sin(omega0*tVals);
sound(testSound);
```

which will play this pitch for 5 seconds. Now, we want to see what happens if we take a mass-on-a-spring with this natural frequency and apply a forcing function with frequency close to this value. The following code assumes a forcing function of frequency 444 Hz. The multiple of ω_0 in front of the forcing function is only for scaling purposes; otherwise the resulting sound would be too quiet.

Code

```
omega = 444*2*pi;

syms ys(t);
[V] = odeToVectorField(diff(ys, 2) + omega0^2*ys == omega0*cos(omega*t));
MS = matlabFunction(V, 'vars', {'t', 'Y'});
soln = ode45(MS, [0,10], [0,0]);

ySound = deval(soln, tVals);
ySound = ySound(1, :);
sound(ySound);
```

A graph of the solution `ySound` can be found in **Normally a reference to a previous figure goes here.**. This exhibits the beating behavior before on a large scale. The sound played during this code also shows the beating or amplitude modulation that can happen in these sorts of solutions. In terms of tuning instruments, these beats are some of the main things musicians will listen for to know if their instrument is close to the right pitch, but just slightly off.

Now suppose $\omega_0 = \omega$. We cannot try the solution $A\cos(\omega t)$ and then use the method of undetermined coefficients, since we notice that $\cos(\omega t)$ solves the associated homogeneous equation. Therefore, we try $x_p = At\cos(\omega t) + Bt\sin(\omega t)$. This time we need the sine term, since the second derivative of $t\cos(\omega t)$ contains sines. We write the equation

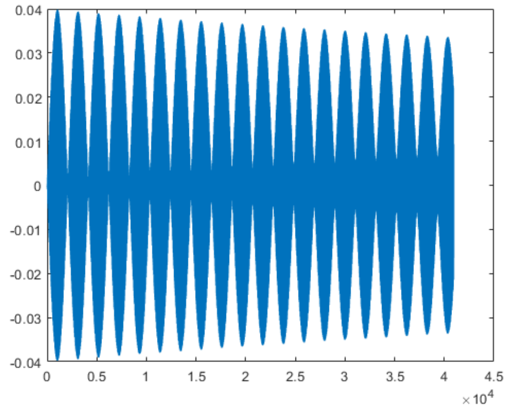


Figure 46: Plot of `ySound` illustrating the beating behavior of interacting sound waves.

$$x'' + \omega^2 x = \frac{F_0}{m} \cos(\omega t).$$

Plugging x_p into the left-hand side we get

$$2B\omega \cos(\omega t) - 2A\omega \sin(\omega t) = \frac{F_0}{m} \cos(\omega t).$$

Hence $A = 0$ and $B = \frac{F_0}{2m\omega}$. Our particular solution is $\frac{F_0}{2m\omega}\, t \sin(\omega t)$ and our general solution is

$$x = C_1 \cos(\omega t) + C_2 \sin(\omega t) + \frac{F_0}{2m\omega}\, t \sin(\omega t).$$

The important term is the last one (the particular solution we found). This term grows without bound as $t \to \infty$. In fact it oscillates between $\frac{F_0 t}{2m\omega}$ and $\frac{-F_0 t}{2m\omega}$. The first two terms only oscillate between $\pm\sqrt{C_1^2 + C_2^2}$, which becomes smaller and smaller in proportion to the oscillations of the last term as t gets larger. In **Normally a reference to a previous figure goes here.** we see the graph with $C_1 = C_2 = 0$, $F_0 = 2$, $m = 1$, $\omega = \pi$.

By forcing the system in just the right frequency we produce very wild oscillations. This kind of behavior is called *resonance* or perhaps *pure resonance*. Sometimes resonance is desired. For example, remember when as a kid you could start swinging by just moving back and forth on the swing seat in the "correct frequency"? You were trying to achieve resonance. The force of each one of your moves was small, but after a while it produced large swings.

On the other hand resonance can be destructive. In an earthquake some buildings collapse while others may be relatively undamaged. This is due to different buildings having different resonance frequencies. So figuring out the resonance frequency can be very important.

A common (but wrong) example of destructive force of resonance is the Tacoma Narrows bridge failure. It turns out there was a different phenomenon at play[1].

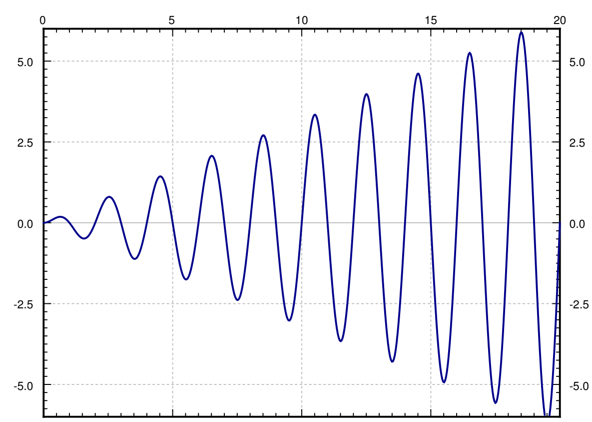


Figure 47: Graph of $\frac{1}{\pi} t \sin(\pi t)$.

[1] K. Billah and R. Scanlan, *Resonance, Tacoma Narrows Bridge Failure, and Undergraduate Physics Textbooks*, American Journal of Physics, 59(2), 1991, 118–124, `http://www.ketchum.org/billah/Billah-Scanlan.pdf`

Damped forced motion and practical resonance

In real life things are not as simple as they were above. There is, of course, some damping. Our equation becomes

$$mx'' + \gamma x' + kx = F_0 \cos(\omega t), \tag{25}$$

for some $\gamma > 0$. We solved the homogeneous problem before. We let

$$p = \frac{\gamma}{2m}, \qquad \omega_0 = \sqrt{\frac{k}{m}}.$$

We replace equation (25) with

$$x'' + 2px' + \omega_0^2 x = \frac{F_0}{m}\cos(\omega t).$$

The roots of the characteristic equation of the associated homogeneous problem are $r_1, r_2 = -p \pm \sqrt{p^2 - \omega_0^2}$. The form of the general solution of the associated homogeneous equation depends on the sign of $p^2 - \omega_0^2$, or equivalently on the sign of $\gamma^2 - 4km$, as before:

$$x_c = \begin{cases} C_1 e^{r_1 t} + C_2 e^{r_2 t} & \text{if } \gamma^2 > 4km, \\ C_1 e^{-pt} + C_2 t e^{-pt} & \text{if } \gamma^2 = 4km, \\ e^{-pt}\big(C_1 \cos(\omega_1 t) + C_2 \sin(\omega_1 t)\big) & \text{if } \gamma^2 < 4km, \end{cases}$$

where $\omega_1 = \sqrt{\omega_0^2 - p^2}$. In any case, we see that $x_c(t) \to 0$ as $t \to \infty$.

Let us find a particular solution. There can be no conflicts when trying to solve for the undetermined coefficients by trying $x_p = A\cos(\omega t) + B\sin(\omega t)$, because the solution to the homogeneous problem will always have exponential factors (since we have damping) and so there is no ω where this will exactly match the form of the homogeneous solution. Let us plug in and solve for A and B. We get (the tedious details are left to reader)

$$\big((\omega_0^2 - \omega^2)B - 2\omega p A\big)\sin(\omega t) + \big((\omega_0^2 - \omega^2)A + 2\omega p B\big)\cos(\omega t) = \frac{F_0}{m}\cos(\omega t).$$

We solve for A and B:

$$A = \frac{(\omega_0^2 - \omega^2)F_0}{m(2\omega p)^2 + m(\omega_0^2 - \omega^2)^2},$$
$$B = \frac{2\omega p F_0}{m(2\omega p)^2 + m(\omega_0^2 - \omega^2)^2}.$$

We also compute $C = \sqrt{A^2 + B^2}$ to be

$$C = \frac{F_0}{m\sqrt{(2\omega p)^2 + (\omega_0^2 - \omega^2)^2}}.$$

Thus our particular solution is

$$x_p = \frac{(\omega_0^2 - \omega^2)F_0}{m(2\omega p)^2 + m(\omega_0^2 - \omega^2)^2}\cos(\omega t) + \frac{2\omega p F_0}{m(2\omega p)^2 + m(\omega_0^2 - \omega^2)^2}\sin(\omega t).$$

Or in the alternative notation we have amplitude C and phase shift δ where (if $\omega \neq \omega_0$)

$$\tan\delta = \frac{B}{A} = \frac{2\omega p}{\omega_0^2 - \omega^2}.$$

Hence,

$$x_p = \frac{F_0}{m\sqrt{(2\omega p)^2 + (\omega_0^2 - \omega^2)^2}}\cos(\omega t - \delta).$$

If $\omega = \omega_0$, then $A = 0$, $B = C = \dfrac{F_0}{2m\omega p}$, and $\delta = \pi/2$.

For reasons we will explain in a moment, we call x_c the *transient solution* and denote it by x_{tr}. We call the x_p from above the *steady periodic solution* and denote it by x_{sp}. The general solution is

$$x = x_c + x_p = x_{tr} + x_{sp}.$$

The transient solution $x_c = x_{tr}$ goes to zero as $t \to \infty$, as all the terms involve an exponential with a negative exponent. So for large t, the effect of x_{tr} is negligible and we see essentially only x_{sp}. Hence the name *transient*. Notice that x_{sp} involves no arbitrary constants, and the initial conditions only affect x_{tr}. Thus, the effect of the initial conditions is negligible after some period of time. We might as well focus on the steady periodic solution and ignore the transient solution. See **Normally a reference to a previous figure goes here.** for a graph given several different initial conditions.

The speed at which x_{tr} goes to zero depends on p (and hence γ). The bigger p is (the bigger γ is), the "faster" x_{tr} becomes negligible. So the smaller the damping, the longer the "transient region". This is consistent with the observation that when $\gamma = 0$, the initial conditions affect the behavior for all time (i.e. an infinite "transient region").

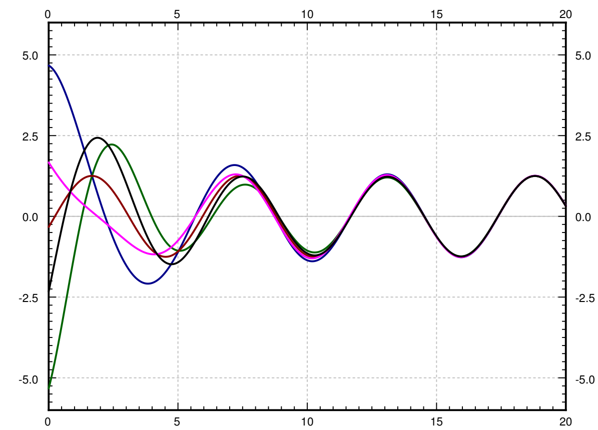


Figure 48: Solutions with different initial conditions for parameters $k = 1$, $m = 1$, $F_0 = 1$, $\gamma = 0.7$, and $\omega = 1.1$.

Let us describe what we mean by resonance when damping is present. Since there were no conflicts when solving with undetermined coefficient, there is no term that goes to infinity. We look instead at the maximum value of the amplitude of the steady periodic solution. Let C be the amplitude of x_{sp}. If we plot C as a function of ω (with all other parameters fixed), we can find its maximum. We call the ω that achieves this maximum the *practical resonance frequency*. We call the maximal amplitude $C(\omega)$ the *practical resonance amplitude*. Thus when damping is present we talk of *practical resonance* rather than pure resonance. A sample plot for three different values of γ is given in **Normally a reference to a previous figure goes here.**. As you can see the practical resonance amplitude grows as damping gets smaller, and practical resonance can disappear altogether when damping is large.

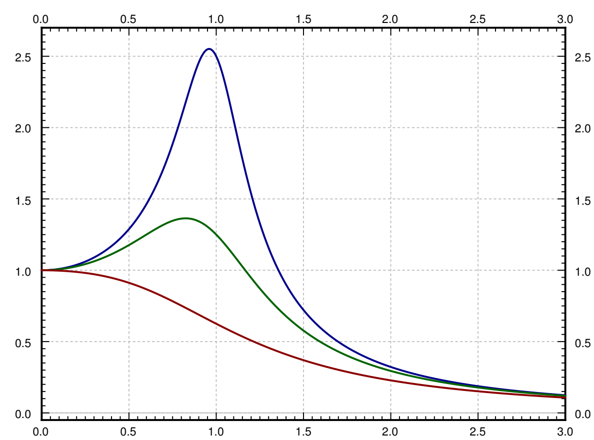


Figure 49: Graph of $C(\omega)$ showing practical resonance with parameters $k = 1$, $m = 1$, $F_0 = 1$. The top line is with $\gamma = 0.4$, the middle line with $\gamma = 0.8$, and the bottom line with $\gamma = 1.6$.

The main takeaways from **Normally a reference to a previous figure goes here.** is that the amplitude can be larger than 1, which is the idea of resonance in this case. Based on Hooke's law, we know that a constant force of magnitude F_0 will stretch (or compress) a spring with constant k a length of F_0/k. If we take $F_0 = 1$ and $k = 1$, as is done in **Normally a reference to a previous figure goes here.**, then the resulting magnitude should be 1. However, if we don't use a constant force of magnitude F_0, but instead use an oscillatory force with frequency ω of the form $F(t) = F_0 \cos(\omega t)$, we get an amplitude of $C(\omega)$. This graph indicates how the forcing frequency changes the amplitude of the resulting oscillation. Since the amplitude "should" be 1 based on F_0/k, if $C(\omega) > 1$, then the frequency chosen is causing an increase in the amplitude, which is the idea of practical resonance.

To find the maximum, or determine if there is a maximum, we need to find the derivative $C'(\omega)$. Computation shows

$$C'(\omega) = \frac{-2\omega(2p^2 + \omega^2 - \omega_0^2)F_0}{m\big((2\omega p)^2 + (\omega_0^2 - \omega^2)^2\big)^{3/2}}.$$

This is zero either when $\omega = 0$ or when $2p^2 + \omega^2 - \omega_0^2 = 0$. In other words, $C'(\omega) = 0$ when

$$\omega = \sqrt{\omega_0^2 - 2p^2} \quad \text{or} \quad \omega = 0.$$

If $\omega_0^2 - 2p^2$ is positive, then there is a positive value of ω, namely $\omega = \sqrt{\omega_0^2 - 2p^2}$ where the amplitude attains a maximum value. Since we know that the amplitude is $F_0/(m\omega_0^2)$ or F_0/k when $\omega = 0$, the maximum will be larger than this. As described above, this value, F_0/k is the expected amplitude, that is, the amplitude you would get with no oscillation, so that if the amplitude is larger than this for some value of ω, this means that the oscillation at frequency ω is resonating with the system to create a larger oscillation. This is the idea of *practical resonance*. It is practical because there is damping, so the situation is more physically relevant (to contrast with pure resonance), and still results in larger amplitudes of oscillation.

Our previous work indicates that a system will exhibit practical resonance for some values of ω whenever $\omega_0^2 - 2p^2$ is positive, and the frequency where the amplitude hits the maximum value is at $\sqrt{\omega_0^2 - 2p^2}$. This follows by the first derivative test for example as then $C'(\omega) > 0$ for small ω in this case. If on the other hand $\omega_0^2 - 2p^2$ is not positive, then $C(\omega)$ achieves its maximum at $\omega = 0$, and there is no practical resonance since we assume $\omega > 0$ in our system. In this case the amplitude gets larger as the forcing frequency gets smaller.

If practical resonance occurs, the peak frequency is smaller than ω_0. As the damping γ (and hence p) becomes smaller, the peak practical resonance frequency goes to ω_0. So when damping is very small, ω_0 is a good estimate of the peak practical resonance frequency. This behavior agrees with the observation that when $\gamma = 0$, then ω_0 is the resonance frequency.

Another interesting observation to make is that when $\omega \to \infty$, then $C \to 0$. This means that if the forcing frequency gets too high it does not manage to get the mass moving in the mass-spring system. This is quite reasonable intuitively. If we wiggle back and forth really fast while sitting on a swing, we will not get it moving at all, no matter how forceful. Fast vibrations just cancel each other out before the mass has any chance of responding by moving one way or the other.

The behavior is more complicated if the forcing function is not an exact cosine wave, but for example a square wave. A general periodic function will be the sum (superposition) of many cosine waves of different frequencies. The reader is encouraged to come back to this section once we have learned about the ideas of Fourier series.

higherOrder/practice/forcedOscillations-practice1.tex

Practice for Forced Oscillations

Why?

Exercise 341 *Write* $\cos(3x) - \cos(2x)$ *as a product of two sine functions.*

Exercise 342 *Write* $\cos(5x) - \cos(3x)$ *as a product of two sine functions.*

Exercise 343 *Write* $\cos(3x) - \cos(\pi x)$ *as a product of two sine functions.*

Exercise 344 *Derive a formula for* x_{sp} *if the equation is* $mx'' + \gamma x' + kx = F_0 \sin(\omega t)$. *Assume* $\gamma > 0$.

Exercise 345 *Derive a formula for* x_{sp} *if the equation is* $mx'' + \gamma x' + kx = F_0 \cos(\omega t) + F_1 \cos(3\omega t)$. *Assume* $\gamma > 0$.

Exercise 346 *Derive a formula for* x_{sp} *for* $mx'' + \gamma x' + kx = F_0 \cos(\omega t) + A$, *where* A *is some constant. Assume* $\gamma > 0$.

Exercise 347 *Take* $mx'' + \gamma x' + kx = F_0 \cos(\omega t)$. *Fix* $m > 0$, $k > 0$, *and* $F_0 > 0$. *Consider the function* $C(\omega)$. *For what values of* γ *(solve in terms of* m, k, *and* F_0*) will there be no practical resonance (that is, for what values of* γ *is there no maximum of* $C(\omega)$ *for* $\omega > 0$*)?*

Exercise 348 *Take* $mx'' + \gamma x' + kx = F_0 \cos(\omega t)$. *Fix* $\gamma > 0$, $k > 0$, *and* $F_0 > 0$. *Consider the function* $C(\omega)$. *For what values of* m *(solve in terms of* γ, k, *and* F_0*) will there be no practical resonance (that is, for what values of* m *is there no maximum of* $C(\omega)$ *for* $\omega > 0$*)?*

Exercise 349 *A mass of* $4\,\mathrm{kg}$ *on a spring with* $k = 4\,\mathrm{N/m}$ *and a damping constant* $c = 1\,\mathrm{Ns/m}$. *Suppose that* $F_0 = 2\,\mathrm{N}$. *Using forcing function* $F_0 \cos(\omega t)$, *find the* ω *that causes the maximum amount of practical resonance and find the amplitude.*

Exercise 350 *An infant is bouncing in a spring chair. The infant has a mass of* $8\,\mathrm{kg}$, *and the chair functions as a spring with spring constant* $72\,\mathrm{N/m}$. *The bouncing of the infant applies a force of the form* $6\cos(\omega t)$ *for some frequency* ω. *Assume that the infant starts at rest at the equilibrium position of the chair.*

a) *If there is no dampening coefficient, what frequency would the infant need to force at in order to generate pure resonance?*

b) *Assume that the chair is built with a dampener with coefficient* $16\,\mathrm{Ns/m}$. *Set up an initial value problem for this situation if the child behaves in the same way.*

Author(s): Matthew Charnley and Jason Nowell

c) *Solve this initial value problem.*

d) *There are several options for chairs you can buy. There is the one with dampening coefficient 16* Ns/m*, one with 1* Ns/m*, and one with 30* Ns/m*. Which of these would be most 'fun' for the infant? How do you know?*

Exercise 351 *A water tower in an earthquake acts as a mass-spring system. Assume that the container on top is full and the water does not move around. The container then acts as the mass and the support acts as the spring, where the induced vibrations are horizontal. The container with water has a mass of* $m = 10{,}000\,\mathrm{kg}$*. It takes a force of 1000 newtons to displace the container 1 meter. For simplicity assume no friction. When the earthquake hits the water tower is at rest (it is not moving). The earthquake induces an external force* $F(t) = mA\omega^2 \cos(\omega t)$*.*

a) *What is the natural frequency of the water tower?*

b) *If* ω *is not the natural frequency, find a formula for the maximal amplitude of the resulting oscillations of the water container (the maximal deviation from the rest position). The motion will be a high frequency wave modulated by a low frequency wave, so simply find the constant in front of the sines.*

c) *Suppose* $A = 1$ *and an earthquake with frequency 0.5 cycles per second comes. What is the amplitude of the oscillations? Suppose that if the water tower moves more than 1.5 meter from the rest position, the tower collapses. Will the tower collapse?*

Exercise 352 *Suppose there is no damping in a mass and spring system with* $m = 5$, $k = 20$, *and* $F_0 = 5$*. Suppose* ω *is chosen to be precisely the resonance frequency.*

a) *Find* ω*.*

b) *Find the amplitude of the oscillations at time* $t = 100$*, given the system is at rest at* $t = 0$*.*

Exercise 353 *Assume that a 2 kg mass is attached to a spring that is acted on by a forcing function* $F(t) = 5\cos(2t)$*. Assume that there is no dampening on the spring.*

a) *What should the spring constant* k *be in order for this system to exhibit pure resonance?*

b) *If we wanted the system to exhibit practical resonance instead, what do or can we change about it to get this?*

c) *Assume that we set* k *to be the value determined in (a), and that the rest of the problem is situated so that the system exhibits practical resonance. What would we expect to see for the amplitude of the solution? This should be a generic comment, not a specific value.*

Exercise 354 *Assume that we have a mass-on-a-spring system defined by the equation*

$$3y'' + 2y' + 18y = 4\cos(5t).$$

a) *Identify the mass, dampening coefficient, and spring constant for the system.*

b) *Use the entire equation to find the natural frequency, forcing frequency, and quasi-frequency of this oscillation.*

c) *Two of these frequencies will show up in the general solution to this problem. Which are they, and in which part (transient, steady-periodic) do they appear?*

d) *Find the general solution of this problem.*

Exercise 355 *A circuit is built with an L Henry inductor, and R Ohm resistor, and a C Farad capacitor. All of the units are correct, but you do not know any of their values. To study this circuit, you apply an external voltage source of $F(t) = 4\cos\left(\frac{1}{2}t\right)$, and the circuit starts with no initial charge or current.*

a) *Write an initial value problem to model this situation.*

b) *Your friend (who knows more about this circuit than you do) takes a reading from this circuit after it is running and says "The amplitude of the charge oscillation is greater than 100 coulombs, which means this circuit is exhibiting practical resonance." There are* **three** *facts that you can learn about this circuit from the statement here that will tell you about the values of L, R, and C.*

   (a) *This statement seems to imply that the expected amplitude of the oscillation is 100 coulombs. What does this mean about the value of C?*

   (b) *Your friend says that this circuit is in practical resonance. What does this tell you about the value of R in this case?*

   (c) *Finally, being in practical resonance says something about how the forcing frequency compares to the natural frequency of this system. What is that, and how does it relate to the value of L?*

c) *What is the frequency of the steady-periodic oscillation that your friend mentioned above?*

Higher order linear ODEs

We discuss Higher order linear ODEs

In this section, we will briefly study higher order equations. Equations appearing in applications tend to be second order. Higher order equations do appear from time to time, but generally the world around us is “second order”.

The basic results about linear ODEs of higher order are essentially the same as for second order equations, with 2 replaced by n. The important concept of linear independence is somewhat more complicated when more than two functions are involved. For higher order constant coefficient ODEs, the methods developed are also somewhat harder to apply, but we will not dwell on these complications. It is also possible to use the methods for systems of linear equations from chapter **??** to solve higher order constant coefficient equations.

Let us start with a general homogeneous linear equation

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_1(x)y' + p_0(x)y = 0. \tag{26}$$

Theorem 11. *Superposition Suppose y_1, y_2, …, y_n are solutions of the homogeneous equation (26). Then*

$$y(x) = C_1y_1(x) + C_2y_2(x) + \cdots + C_ny_n(x)$$

also solves (26) for arbitrary constants $C_1, C_2, \ldots, C_n$.

In other words, a *linear combination* of solutions to (26) is also a solution to (26). We also have the existence and uniqueness theorem for nonhomogeneous linear equations.

Theorem 12. *Existence and uniqueness Suppose p_0 through p_{n-1}, and f are continuous functions on some interval I, a is a number in I, and $b_0, b_1, \ldots, b_{n-1}$ are constants. The equation*

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_1(x)y' + p_0(x)y = f(x)$$

has exactly one solution $y(x)$ defined on the same interval I satisfying the initial conditions

$$y(a) = b_0, \quad y'(a) = b_1, \quad \ldots, \quad y^{(n-1)}(a) = b_{n-1}.$$

Linear independence

When we had two functions y_1 and y_2 we said they were linearly independent if one was not the multiple of the other. Same idea holds for n functions. In this case, it is easier to state as follows. The functions y_1, y_2, …, y_n are *linearly independent* if the equation

$$c_1y_1 + c_2y_2 + \cdots + c_ny_n = 0$$

has only the trivial solution $c_1 = c_2 = \cdots = c_n = 0$, where the equation must hold for all x. If we can solve equation with some constants where for example $c_1 \neq 0$, then we can solve for y_1 as a linear combination of the others. If the functions are not linearly independent, they are *linearly dependent*.

Example 61. *Show that e^x, e^{2x}, e^{3x} are linearly independent.*

Solution: Let us give several ways to show this fact. Many textbooks (including and) introduce Wronskians for higher order equations, but it is harder to analyze them without tools from linear algebra (see Chapter **??**). Once there are more than two functions involved, there is not a nice, simple formula for the Wronskian (like $y_1'y_2 - y_2'y_1$ for two functions) and linear algebra is required to analyze what is happening here. Instead, we will take a slightly different and more improvized approach to see why these functions are linearly independent.

Let us write down

$$c_1e^x + c_2e^{2x} + c_3e^{3x} = 0.$$

Learning outcomes: Find the general solution to a linear, constant coefficient, homogeneous differential equation of higher order Solve non-homogeneous higher order equations using the method of undetermined coefficients.

Author(s): Matthew Charnley and Jason Nowell

We use rules of exponentials and write $z = e^x$. Hence $z^2 = e^{2x}$ and $z^3 = e^{3x}$. Then we have

$$c_1 z + c_2 z^2 + c_3 z^3 = 0.$$

The left-hand side is a third degree polynomial in z. It is either identically zero, or it has at most 3 zeros. Therefore, it is identically zero, $c_1 = c_2 = c_3 = 0$, and the functions are linearly independent.

Let us try another way. As before we write

$$c_1 e^x + c_2 e^{2x} + c_3 e^{3x} = 0.$$

This equation has to hold for all x. We divide through by e^{3x} to get

$$c_1 e^{-2x} + c_2 e^{-x} + c_3 = 0.$$

As the equation is true for all x, let $x \to \infty$. After taking the limit we see that $c_3 = 0$. Hence our equation becomes

$$c_1 e^x + c_2 e^{2x} = 0.$$

Rinse, repeat!

How about yet another way. We again write

$$c_1 e^x + c_2 e^{2x} + c_3 e^{3x} = 0.$$

We can evaluate the equation and its derivatives at different values of x to obtain equations for c_1, c_2, and c_3. Let us first divide by e^x for simplicity.

$$c_1 + c_2 e^x + c_3 e^{2x} = 0.$$

We set $x = 0$ to get the equation $c_1 + c_2 + c_3 = 0$. Now differentiate both sides

$$c_2 e^x + 2c_3 e^{2x} = 0.$$

We set $x = 0$ to get $c_2 + 2c_3 = 0$. We divide by e^x again and differentiate to get $2c_3 e^x = 0$. It is clear that c_3 is zero. Then c_2 must be zero as $c_2 = -2c_3$, and c_1 must be zero because $c_1 + c_2 + c_3 = 0$.

There is no one best way to do it. All of these methods are perfectly valid. The important thing is to understand why the functions are linearly independent.

Exercise 356 *Here is the linear algebra method for after reading through that chapter. Let $y_1 = e^x$, $y_2 = e^{2x}$ and $y_3 = e^{3x}$. Verify that*

$$\begin{bmatrix} y_1(0) \\ y_1'(0) \\ y_1''(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} y_2(0) \\ y_2'(0) \\ y_2''(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \qquad \begin{bmatrix} y_3(0) \\ y_3'(0) \\ y_3''(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix}$$

and use that to determine that these functions are linearly independent by showing that

$$\det \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix} = 2 \neq 0$$

so that this matrix is invertible.

Example 62. *On the other hand, the functions e^x, e^{-x}, and $\cosh x$ are linearly dependent. Simply apply definition of the hyperbolic cosine:*

$$\cosh x = \frac{e^x + e^{-x}}{2} \qquad \text{or} \qquad 2\cosh x - e^x - e^{-x} = 0.$$

This second form here is a linear combination (coefficients 2, -1, and -1) of the three functions that adds to zero.

Constant coefficient higher order ODEs

When we have a higher order constant coefficient homogeneous linear equation, the song and dance is exactly the same as it was for second order. We just need to find more solutions. If the equation is n^{th} order, we need to find n linearly independent solutions. It is best seen by example.

Example 63. *Find the general solution to*

$$y''' - 3y'' - y' + 3y = 0. \tag{27}$$

Solution: Try: $y = e^{rx}$. We plug in and get

$$\underbrace{r^3 e^{rx}}_{y'''} - 3\underbrace{r^2 e^{rx}}_{y''} - \underbrace{r e^{rx}}_{y'} + 3\underbrace{e^{rx}}_{y} = 0.$$

We divide through by e^{rx}. Then

$$r^3 - 3r^2 - r + 3 = 0.$$

The trick now is to find the roots. There is a formula for the roots of degree 3 and 4 polynomials but it is very complicated. There is no formula for higher degree polynomials. That does not mean that the roots do not exist. There are always n roots for an n^{th} degree polynomial. They may be repeated and they may be complex. Computers are pretty good at finding roots approximately for reasonable size polynomials.

A good place to start is to plot the polynomial and check where it is zero. We can also simply try plugging in. We just start plugging in numbers $r = -2, -1, 0, 1, 2, \ldots$ and see if we get a hit (we can also try complex numbers). Even if we do not get a hit, we may get an indication of where the root is. For example, we plug $r = -2$ into our polynomial and get -15; we plug in $r = 0$ and get 3. That means there is a root between $r = -2$ and $r = 0$, because the sign changed. If we find one root, say r_1, then we know $(r - r_1)$ is a factor of our polynomial. Polynomial long division can then be used.

Another technique for guessing roots of polynomials is the Rational Roots Theorem, which says that any rational root of the polynomial must be of the form p/q where p divides the constant term of the polynomial and q divides the leading term, provided neither of them are zero. For more information on this see § **??**. In this case, we would know that p must divide 3, and q must divide 1. Therefore, the only possible options here are ± 1 and ± 3. These would be good places to start to look for rational roots.

A good strategy is to begin with $r = 0$, 1, or -1. These are easy to compute. Our polynomial has two such roots, $r_1 = -1$ and $r_2 = 1$. There should be 3 roots and the last root is reasonably easy to find. The constant term in a monic[1] polynomial such as this is the multiple of the negations of all the roots because $r^3 - 3r^2 - r + 3 = (r - r_1)(r - r_2)(r - r_3)$. So

$$3 = (-r_1)(-r_2)(-r_3) = (1)(-1)(-r_3) = r_3.$$

You should check that $r_3 = 3$ really is a root. Hence e^{-x}, e^{x} and e^{3x} are solutions to (27). They are linearly independent as can easily be checked, and there are 3 of them, which happens to be exactly the number we need. So the general solution is

$$y = C_1 e^{-x} + C_2 e^{x} + C_3 e^{3x}.$$

Another possible way to work out this general solution is by factoring the original polynomial. Since we want to solve

$$r^3 - 3r^2 - r + 3 = 0,$$

we can rewrite the polynomial as

$$r^2(r-3) - 1(r-3) = 0$$

which factors as

$$(r^2 - 1)(r - 3) = 0.$$

Finally, using difference of two squares on the first factor gives

$$(r-1)(r+1)(r-3) = 0.$$

, This gives roots of 1, -1, and 3, and so the same general solution as above.

[1]The word monic means that the coefficient of the top degree r^d, in our case r^3, is 1.

Suppose we were given some initial conditions $y(0) = 1$, $y'(0) = 2$, and $y''(0) = 3$. Then

$$\begin{aligned}
1 = y(0) &= C_1 + C_2 + C_3, \\
2 = y'(0) &= -C_1 + C_2 + 3C_3, \\
3 = y''(0) &= C_1 + C_2 + 9C_3.
\end{aligned}$$

It is possible to find the solution by high school algebra, but it would be a pain. The sensible way to solve a system of equations such as this is to use matrix algebra, see § **??** or Chapter **??**. For now we note that the solution is $C_1 = -1/4$, $C_2 = 1$, and $C_3 = 1/4$. The specific solution to the ODE is

$$y = \frac{-1}{4}\, e^{-x} + e^{x} + \frac{1}{4}\, e^{3x}.$$

Next, suppose that we have real roots, but they are repeated. Let us say we have a root r repeated k times. In the spirit of the second order solution, and for the same reasons, we have the solutions

$$e^{rx}, \quad xe^{rx}, \quad x^2e^{rx}, \quad \ldots, \quad x^{k-1}e^{rx}.$$

We take a linear combination of these solutions to find the general solution.

Example 64. *Solve*

$$y^{(4)} - 3y''' + 3y'' - y' = 0.$$

Solution: We note that the characteristic equation is

$$r^4 - 3r^3 + 3r^2 - r = 0.$$

By inspection we note that $r^4 - 3r^3 + 3r^2 - r = r(r-1)^3$. Hence the roots given with multiplicity are $r = 0, 1, 1, 1$. Thus the general solution is

$$y = \underbrace{(C_1 + C_2x + C_3x^2)\, e^{x}}_{\text{terms coming from } r=1} + \underbrace{C_4}_{\text{from } r=0}.$$

Example 65. *Find the general solution of*

$$y''' + 2y'' - 5y' - 6y = 0$$

Solution: The characteristic equation for this example is

$$r^3 + 2r^2 - 5r - 6 = 0.$$

There is no convenient factoring by grouping or other quick formula to get to the roots here. The best hope we have is to try to guess the roots and see if we come up with anything. Once we get one root, we'll be able to factor a term out and get down to a quadratic equation, where the quadratic formula will give us the other two roots.

The properties of polynomials tell us that all rational roots of this polynomial must be factors of $\frac{-6}{1}$ or -6. Thus, the options are ± 1, ± 2, and ± 3. At this point, the best bet is to start guessing and see if we can find one. Let's start with 1. Plugging this into the polynomial gives

$$1^3 + 2(1)^2 - 5(1) - 6 = -8 \neq 0.$$

Trying -1 next, we get

$$(-1)^3 + 2(-1)^2 - 5(-1) - 6 = -1 + 2 + 5 - 6 = 0.$$

Therefore, -1 is as root, and so $(r+1)$ is a factor of this polynomial.

We can then use synthetic (or long) division to see that

$$r^3 + 2r^2 - 5r - 6 = (r+1)(r^2 + r - 6).$$

For the quadratic, we can either use the quadratic formula, or just recognize that this factors as $(r-2)(r+3)$ to get that the characteric equation factors as

$$(r+1)(r-2)(r+3) = 0.$$

Therefore, the roots are -1, 2 and -3, so that the general solution to the differential equation is

$$y(x) = C_1 e^{-x} + C_2 e^{2x} + C_3 e^{-3x}.$$

For more information on synthetic division and finding roots of polynomials, see **Normally a reference to an Appendix goes here.**.

The case of complex roots is similar to second order equations. Complex roots always come in pairs $r = \alpha \pm i\beta$. Suppose we have two such complex roots, each repeated k times. The corresponding solution is

$$(C_0 + C_1 x + \cdots + C_{k-1}x^{k-1})\, e^{\alpha x}\cos(\beta x) + (D_0 + D_1 x + \cdots + D_{k-1}x^{k-1})\, e^{\alpha x}\sin(\beta x).$$

where $C_0, \ldots, C_{k-1}, D_0, \ldots, D_{k-1}$ are arbitrary constants.

Example 66. *Solve*

$$y^{(4)} - 4y''' + 8y'' - 8y' + 4y = 0.$$

Solution: The characteristic equation is

$$\begin{aligned} r^4 - 4r^3 + 8r^2 - 8r + 4 &= 0, \\ (r^2 - 2r + 2)^2 &= 0, \\ \left((r-1)^2 + 1\right)^2 &= 0. \end{aligned}$$

Hence the roots are $1 \pm i$, both with multiplicity 2. Hence the general solution to the ODE is

$$y = (C_1 + C_2 x)\, e^x \cos x + (C_3 + C_4 x)\, e^x \sin x.$$

The way we solved the characteristic equation above is really by guessing or by inspection. It is not so easy in general. We could also have asked a computer or an advanced calculator for the roots.

Non-Homogeneous Equations

Just like for second order equation, we can solve higher order non-homogeneous equations. The theory is the same; if we can find any single solution to the non-homogeneous problem, then the general solution of the non-homogeneous problem is this single solution plus the general solution to the corresponding homogeneous problem. The trick comes down to finding this single solution, and undetermined coefficients is the main method here.

In using undetermined coefficients, the guesses we want to make are the same as for second order equations. The only way it really gets more complicated is that now it is possible for any exponential or trigonometric function to be a solution to the homogeneous problem, and so more things will need to be multiplied by x in order to get the appropriate guess for the non-homogeneous solution.

Example 67. *Find the general solution to*

$$y''' + 2y'' - 5y' - 6y = 3e^{2x} + e^{4x}.$$

Solution: We found the general solution of the homogeneous problem in **Normally a reference to a previous example goes here.**, which is

$$y(x) = C_1 e^{-x} + C_2 e^{2x} + C_3 e^{-3x}.$$

Now, to solve the non-homogeneous problem, we use the method of undetermined coefficients. Since the non-homogeneous part of the equation has terms of the form e^{2x} and e^{4x}, we would want to guess

$$y_p(x) = Ae^{2x} + Be^{4x}.$$

However, e^{2x} solves the homogeneous problem, so we need to multiply it by x, making our actual guess become

$$y_p(x) = Axe^{2x} + Be^{4x}.$$

In order to plug this in, we need to take three derivatives of this guess, which are

$$\begin{aligned} y_p(x) &= Axe^{2x} + Be^{4x} \\ y_p'(x) &= Ae^{2x} + 2Axe^{2x} + 4Be^{4x} \\ y_p''(x) &= 4Ae^{2x} + 4Axe^{2x} + 16Be^{4x} \\ y_p'''(x) &= 12Ae^{2x} + 8Axe^{2x} + 64Be^{4x} \end{aligned}.$$

By putting this into the non-homogeneous equation we want to solve, we get

$$\begin{aligned} &(12Ae^{2x} + 8Axe^{2x} + 64Be^{4x}) + 2(4Ae^{2x} + 4Axe^{2x} + 16Be^{4x}) \\ &\quad - 5(Ae^{2x} + 2Axe^{2x} + 4Be^{4x}) - 6(Axe^{2x} + Be^{4x}) = 3e^{2x} + e^{4x} \end{aligned}.$$

Simplifying the left hand side of this expression gives

$$15Ae^{2x} + 70Be^{4x} = 3e^{2x} + e^{4x}.$$

To satisfy this equation, we want to set $A = \frac{1}{5}$ and $B = \frac{1}{70}$. Therefore, the general solution to the non-homogeneous problem is

$$y(x) = C_1e^{-x} + C_2e^{2x} + C_3e^{-3x} + \frac{1}{5}xe^{2x} + \frac{1}{70}e^{4x}.$$

Example 68. *Determine the form of the guess using undetermined coefficients for finding a particular solution of the non-homogeneous problem*

$$y^{(9)} + y^{(8)} - 2y^{(5)} - 2y^{(4)} + y' + y = e^x + 3e^{-x} + \sin(x) + 2x.$$

Solution: To determine the guess, we need to first find the solution to the homogeneous equations. The characteristic equation of the homogeneous equation is

$$r^9 + r^8 - 2r^5 - 2r^4 + r + 1 = 0.$$

We could use the root guessing method for this example, and all rational roots must be ± 1. However, that method is not great for polynomials that are of degree higher than around 3 or 4. So, we'll want to use some other technique to find all of the root.

If we start by grouping pairs of terms, we can rewrite this polynomial as

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

so that it can be rewritten as

$$(r+1)(r^8 - 2r^4 + 1) = 0.$$

The second factor looks a lot like

$$(s-1)^2 = s^2 - 2s + 1$$

if we take $s = r^4$. Since

$$(r^4 - 1) = (r^2+1)(r^2-1) = (r^2+1)(r+1)(r-1)$$

using difference of squares twice. Thus, the entire characteristic equation can be written as

$$(r+1)(r^4-1)^2 = (r+1)[(r^2+1)(r+1)(r-1)]^2 = (r+1)^3(r-1)^2(r^2+1)^2.$$

Therefore, we have a triple root at -1, a double root at 1, and two copies of (r^2+1), which has a root of i, corresponding to solutions $\sin(x)$ and $\cos(x)$. Putting all of this together, the general solution to the homogeneous equation is

$$y_c(x) = (C_1 + C_2x + C_3x^2)e^{-x} + (C_4 + C_5x)e^x + (C_6 + C_7x)\sin(x) + (C_8 + C_9x)\cos(x).$$

This has 9 unknown constants in it, which is expected from the ninth order equation.

Now, we need to figure out the appropriate guess for the non-homogeneous solution. Since the non-homogeneous part of the equation is $e^x + 3e^{-x} + \sin x + 2x$, the base guess would be of the form

$$Ae^x + Be^{-x} + C\sin x + D\cos x + Ex + F$$

because we always need to include both $\sin(x)$ and $\cos(x)$ whenever either of them appear. However, we need to factor in what terms show up in the homogeneous solution. For instance, the e^x term has a term with 1 and x in the homogeneous solution, we need to include the next one up in our guess for the solution to the non-homogeneous problem. Taking this into account for all terms gives the desired guess as

$$y_p(x) = Ax^2e^x + Bx^3e^{-x} + Cx^2\sin(x) + Dx^2\cos(x) + Ex + F.$$

$\lrcorner$

There is also an extension of variation of parameters to higher order equations. However, the fact that there are more terms in the solution means that the form of the expression is much more complicated that for second order, and is not worth looking into or trying to remember. The easier way to handle these situations using variation of parameters is by converting the higher order equation into a first order system and applying the methods there, which will be covered in § **??** and § respectively.

higherOrder/practice/higherOrderODE-practice1.tex

Practice for Higher Order ODEs

Why?

Exercise 357 *Find the general solution for* $y''' - y'' + y' - y = 0$.

Exercise 358 *Find the general solution of* $y^{(5)} - y^{(4)} = 0$.

Exercise 359 *Find the general solution for* $y^{(4)} - 5y''' + 6y'' = 0$.

Exercise 360 *Find the general solution for* $y''' + 2y'' + 2y' = 0$.

Exercise 361 *Suppose the characteristic equation for an ODE is* $(r-1)^2(r-2)^2 = 0$.

a) *Find such a differential equation.*
b) *Find its general solution.*

Exercise 362 *Suppose that a fourth order equation has a solution* $y = 2e^{4x}x\cos x$.

a) *Find such an equation.*
b) *Find the initial conditions that the given solution satisfies.*

Exercise 363 *Suppose that the characteristic equation of a third order differential equation has roots* $\pm 2i$ *and 3.*

a) *What is the characteristic equation?*
b) *Find the corresponding differential equation.*
c) *Find the general solution.*

Exercise 364 *Find the general solution for the equation of* Exercise .

Exercise 365 *Find the general solution of*

$$y^{(4)} - y''' - 5y'' - 23y' - 20y = 0.$$

Author(s): Matthew Charnley and Jason Nowell

Exercise 366 *Find the general solution of*

$$y''' - 6y'' + 13y' - 10y = 4e^x + 5e^{3x} - 20.$$

Exercise 367 *Find the general solution of*

$$y''' - 3y' + 2y = 2e^x - e^{3x}.$$

Exercise 368 *Find the general solution of*

$$y''' + 2y'' + y' + 2y = 3\cos(x) + x.$$

Exercise 369 *Find the general solution of*

$$y^{(4)} + 2y'' + y = 4x\cos(x) - e^{3x} + 1$$

Hint: Remember, the guess needs to make sure that no terms in it solve the homogeneous equation.

Exercise 370 *Show that $y = \cos(2t)$ is a solution to $y^{(4)} + 2y''' + 9y'' + 8y' + 20y = 0$. This tells us something about the factorization of the characteristic polynomial of this DE. Factor the characteristic polynomial completely, and solve the DE.*

Exercise 371 *Consider*

$$y''' - y'' - 8y' + 12y = 0. \tag{28}$$

a) *Show that $y = e^{2t}$ is a solution of (28).*

b) *Find the general solution to (28).*

c) *Solve $y''' - y'' - 8y' + 12y = e^{2t}$.*

Exercise 372 *Let $f(x) = e^x - \cos x$, $g(x) = e^x + \cos x$, and $h(x) = \cos x$. Are $f(x)$, $g(x)$, and $h(x)$ linearly independent? If so, show it, if not, find a linear combination that works.*

Exercise 373 *Let $f(x) = 0$, $g(x) = \cos x$, and $h(x) = \sin x$. Are $f(x)$, $g(x)$, and $h(x)$ linearly independent? If so, show it, if not, find a linear combination that works.*

Exercise 374 *Are e^x, e^{x+1}, e^{2x}, $\sin(x)$ linearly independent? If so, show it, if not find a linear combination that works.*

Exercise 375 *Are x, x^2, and x^4 linearly independent? If so, show it, if not, find a linear combination that works.*

Exercise 376 *Are e^x, xe^x, and x^2e^x linearly independent? If so, show it, if not, find a linear combination that works.*

Exercise 377 *Are $\sin(x)$, x, $x\sin(x)$ linearly independent? If so, show it, if not find a linear combination that works.*

Exercise 378 *Show that $\{e^t, te^t, e^{-t}, te^{-t}\}$ is a linearly independent set.*

Exercise 379 *Solve $1001y''' + 3.2y'' + \pi y' - \sqrt{4}y = 0$, $y(0) = 0$, $y'(0) = 0$, $y''(0) = 0$.*

Exercise 380 *Could $y = t^2 \cos t$ be a solution of a homogeneous DE with constant real coefficients? If so, give the minimum possible order of such a DE, and state which functions must also be solutions. If not, explain why this is impossible.*

Exercise 381 *Find a linear DE with constant real coefficients whose general solution is*

$$y = c_1e^{2t} + c_2e^{-t}\cos(4t) + c_3e^{-t}\sin(2t),$$

or explain why there is no such thing.

Exercise 382 *Find an equation such that $y = xe^{-2x}\sin(3x)$ is a solution.*

Exercise 383 *Find an equation of minimal order such that $y = \cos(x)$, $y = \sin(x)$, $y = e^x$ are solutions.*

Exercise 384 *Find an equation of minimal order such that $y = \cos(x)$, $y = \sin(2x)$, $y = e^{3x}$ are solutions.*

Exercise 385 *Find a homogeneous DE with general solution*

$$y = c_1e^t + c_2e^{-t} + c_3\cos t + c_4\sin t + c_5te^t + c_6te^{-t} + c_7t\cos t + c_8t\sin t.$$

Chapter Review

A Set of Problems for a Chapter Review on Higher Order ODEs

Topics

Exercises

Exercise 386 *Find a homogeneous DE with general solution $y = c_1 + c_2 t + c_3 \cos 2t + c_4 \sin 2t$, or explain why there is no such thing.*

Exercise 387 *This exercise looks at a solution to a non-homogeneous higher-order equation.*

a) *Solve $y^{(4)} - 200y'' + 10000y = 0$.*

b) *Write the correct ansatz to solve $y^{(4)} - 200y'' + 10000y = e^{10t} + \cos(10t)$. Do not actually solve that.*

Author(s): Matthew Charnley and Jason Nowell

Vectors, mappings, and matrices

We discuss Vectors, mappings, and matrices

In real life, there is most often more than one variable. We wish to organize dealing with multiple variables in a consistent manner, and in particular organize dealing with linear equations and linear mappings, as those both rather useful and rather easy to handle. Mathematicians joke that "to an engineer every problem is linear, and everything is a matrix." And well, they (the engineers) are not wrong. Quite often, solving an engineering problem is figuring out the right finite-dimensional linear problem to solve, which is then solved with some matrix manipulation. Most importantly, linear problems are the ones that we know how to solve, and we have many tools to solve them. For engineers, mathematicians, physicists, and anybody in a technical field it is absolutely vital to learn linear algebra.

As motivation, suppose we wish to solve

$$\begin{aligned} x - y &= 2, \\ 2x + y &= 4, \end{aligned}$$

for x and y, that is, find numbers x and y such that the two equations are satisfied. Let us perhaps start by adding the equations together to find

$$x + 2x - y + y = 2 + 4, \qquad \text{or} \qquad 3x = 6.$$

In other words, $x = 2$. Once we have that, we plug in $x = 2$ into the first equation to find $2 - y = 2$, so $y = 0$. OK, that was easy. What is all this fuss about linear equations. Well, try doing this if you have 5000 unknowns[1]. Also, we may have such equations not of just numbers, but of functions and derivatives of functions in differential equations. Clearly we need a more systematic way of doing things. A nice consequence of making things systematic and simpler to write down is that it becomes easier to have computers do the work for us. Computers are rather stupid, they do not think, but are very good at doing lots of repetitive tasks precisely, as long as we figure out a systematic way for them to perform the tasks.

Vectors and operations on vectors

Consider n real numbers as an n-tuple:

$$(x_1, x_2, \ldots, x_n).$$

The set of such n-tuples is the so-called *n-dimensional space*, often denoted by $\mathbb{R}^n$. Sometimes we call this the n-dimensional *euclidean space*[2]. In two dimensions, $\mathbb{R}^2$ is called the *cartesian plane*[3], and in three dimensions, it is the same "3-dimensional space" that is dealt with in multivariable calculus. Each such n-tuple represents a point in the n-dimensional space. For example, the point $(1, 2)$ in the plane $\mathbb{R}^2$ is one unit to the right and two units up from the origin.

When we do algebra with these n-tuples of numbers we call them *vectors*[4]. Mathematicians are keen on separating what is a vector and what is a point of the space or in the plane, and it turns out to be an important distinction, however, for the purposes of linear algebra we can think of everything being represented by a vector. A way to think of a vector, which is especially useful in calculus and differential equations, is an arrow. It is an object that has a *direction* and a *magnitude*. For example, the vector $(1, 2)$ is the arrow from the origin to the point $(1, 2)$ in the plane. The magnitude is the length of the arrow. See **Normally a reference to a previous figure goes here.**. If we think of vectors as arrows, the arrow doesn't always have to start at the origin. If we do move it around, however, it should always keep the same direction and the same magnitude.

As vectors are arrows, when we want to give a name to a vector, we draw a little arrow above it:

$$\vec{x}$$

Learning outcomes: Express n-tuples of numbers as vectors Perform operations on vectors Understand how linear maps on vectors give rise to matrices.

Author(s): Matthew Charnley and Jason Nowell

[1] One of the downsides of making everything look like a linear problem is that the number of variables tends to become huge.

[2] Named after the ancient Greek mathematician Euclid of Alexandria (around 300 BC), possibly the most famous of mathematicians; even small towns often have Euclid Street or Euclid Avenue.

[3] Named after the French mathematician René Descartes (1596–1650). It is "cartesian" as his name in Latin is Renatus Cartesius.

[4] A common notation to distinguish vectors from points is to write $(1, 2)$ for the point and $\langle 1, 2\rangle$ for the vector. We write both as $(1, 2)$.

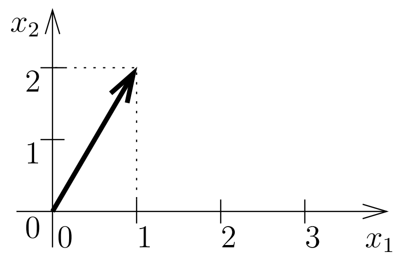


Figure 50: The vector $(1, 2)$ drawn as an arrow from the origin to the point $(1, 2)$.

Another popular notation is $\mathbf{x}$, although we will use the little arrows. It may be easy to write a bold letter in a book, but it is not so easy to write it by hand on paper or on the board. Mathematicians often don't even write the arrows. A mathematician would write x and just remember that x is a vector and not a number. Just like you remember that Bob is your uncle, and you don't have to keep repeating "Uncle Bob" and you can just say "Bob." In this book, however, we will call Bob "Uncle Bob" and write vectors with the little arrows.

The *magnitude* can be computed using Pythagorean theorem. The vector $(1, 2)$ drawn in the figure has magnitude $\sqrt{1^2 + 2^2} = \sqrt{5}$. The magnitude is denoted by $\|\vec{x}\|$, and, in any number of dimensions, it can be computed in the same way:

$$\|\vec{x}\| = \|(x_1, x_2, \ldots, x_n)\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

For reasons that will become clear in the next section, we often write vectors as so-called *column vectors*:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Don't worry. It is just a different way of writing the same thing, and it will be useful later. For example, the vector $(1, 2)$ can be written as

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

The fact that we write arrows above vectors allows us to write several vectors $\vec{x}_1$, $\vec{x}_2$, etc., without confusing these with the components of some other vector $\vec{x}$.

So where is the *algebra* from *linear algebra*? Well, arrows can be added, subtracted, and multiplied by numbers. First we consider *addition*. If we have two arrows, we simply move along one, and then along the other. See **Normally a reference to a previous figure goes here.**.

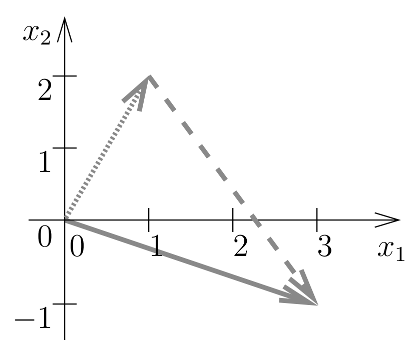


Figure 51: Adding the vectors $(1, 2)$, drawn dotted, and $(2, -3)$, drawn dashed. The result, $(3, -1)$, is drawn as a solid arrow.

It is rather easy to see what it does to the numbers that represent the vectors. Suppose we want to add $(1, 2)$ to $(2, -3)$ as in the figure. So we travel along $(1, 2)$ and then we travel along $(2, -3)$ in the sense of "tip-to-tail" addition that you may have seen in previous classes. What we did was travel one unit right, two units up, and then we travelled two units right, and three units down (the negative three). That means that we ended up at $\bigl(1 + 2, 2 + (-3)\bigr) = (3, -1)$. And that's how addition always works:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}.$$

Subtracting is similar. What $\vec{x} - \vec{y}$ means visually is that we first travel along $\vec{x}$, and then we travel backwards along $\vec{y}$. See **Normally a reference to a previous figure goes here.**. It is like adding $\vec{x} + (-\vec{y})$ where $-\vec{y}$ is the arrow we obtain by erasing the arrow head from one side and drawing it on the other side, that is, we reverse the direction. In terms of the numbers, we simply go backwards in both directions, so we negate both numbers. For example, if $\vec{y}$ is $(-2, 1)$, then $-\vec{y}$ is $(2, -1)$.

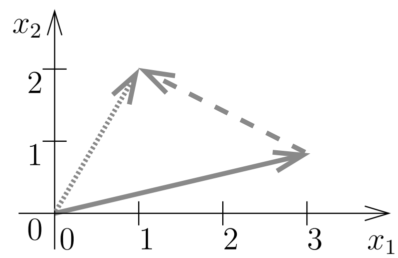


Figure 52: Subtraction, the vector $(1, 2)$, drawn dotted, minus $(-2, 1)$, drawn dashed. The result, $(3, 1)$, is drawn as a solid arrow.

Another intuitive thing to do to a vector is to *scale* it. We represent this by multiplication of a number with a vector. Because of this, when we wish to distinguish between vectors and numbers, we call the numbers *scalars*. For example, suppose we want to travel three times further. If the vector is $(1, 2)$, traveling 3 times further means going 3 units to the right and 6 units up, so we get the vector $(3, 6)$. We just multiply each number in the vector by 3. If α is a number, then

$$\alpha \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}.$$

Scaling (by a positive number) multiplies the magnitude and leaves direction untouched. The magnitude of $(1, 2)$ is $\sqrt{5}$. The magnitude of 3 times $(1, 2)$, that is, $(3, 6)$, is $3\sqrt{5}$.

When the scalar is negative, then when we multiply a vector by it, the vector is not only scaled, but it also switches direction. So multiplying $(1, 2)$ by -3 means we should go 3 times further but in the opposite direction, so 3 units to the left and 6 units down, or in other words, $(-3, -6)$. As we mentioned above, $-\vec{y}$ is a reverse of $\vec{y}$, and this is the same as $(-1)\vec{y}$.

In **Normally a reference to a previous figure goes here.**, you can see a couple of examples of what scaling a vector means visually.

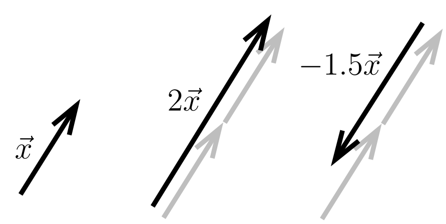


Figure 53: A vector $\vec{x}$, the vector $2\vec{x}$ (same direction, double the magnitude), and the vector $-1.5\vec{x}$ (opposite direction, 1.5 times the magnitude).

We put all of these operations together to work out more complicated expressions. Let us compute a small example:

$$3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 2 \begin{bmatrix} -4 \\ -1 \end{bmatrix} - 3 \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 3(1) + 2(-4) - 3(-2) \\ 3(2) + 2(-1) - 3(2) \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

As we said a vector is a direction and a magnitude. Magnitude is easy to represent, it is just a number. The *direction* is usually given by a vector with magnitude one. We call such a vector a *unit vector*. That is, $\vec{u}$ is a unit vector when $\lVert\vec{u}\rVert = 1$.

For example, the vectors $(1, 0)$, $(1/\sqrt{2}, 1/\sqrt{2})$, and $(0, -1)$ are all unit vectors. To represent the direction of a vector $\vec{x}$, we need to find the unit vector in the same direction. To do so, we simply rescale $\vec{x}$ by the reciprocal of the magnitude, that is $\frac{1}{\lVert\vec{x}\rVert}\vec{x}$, or more concisely $\frac{\vec{x}}{\lVert\vec{x}\rVert}$.

For example, the unit vector in the direction of $(1, 2)$ is the vector

$$\frac{1}{\sqrt{1^2 + 2^2}}(1, 2) = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right).$$

Linear Combinations and Linear Independence

If we have vectors of a given size (say a collection of 3 component vectors), there are only two different operations we can perform on them; adding them together and multiplying them by real numbers. While this may seem fairly limited, we can take a small number of vectors and generate a lot more vectors from them. By putting these operations together, we get the definition of a linear combination.

Definition 11. *Given row or column vectors* $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n$*, a* linear combination *is an expression of the form*

$$\alpha_1\vec{y}_1 + \alpha_2\vec{y}_2 + \cdots + \alpha_n\vec{y}_n,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n$ *are all scalars.*

For example, $3\vec{y}_1 + \vec{y}_2 - 5\vec{y}_3$ is a linear combination of $\vec{y}_1$, $\vec{y}_2$, and $\vec{y}_3$. Another example is that $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ can be written as a linear combination of $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$, and $\begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}$ as

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

Given any collection of vectors $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n$, it is always possible to write $\vec{0}$ as a linear combination of these vectors by choosing $\alpha_i = 0$ for all i. Then we are adding a bunch of zeros together and so get zero. However, this may not be the only way to get zero. Whether or not it is depends heavily on the exact vectors involved.

Definition 12. *We say the vectors* $\vec{y}_1$*,* $\vec{y}_2$*, . . . ,* $\vec{y}_n$ *are* linearly independent *if the only way to pick* $\alpha_1,\ \alpha_2, ..., \alpha_n$ *to satisfy*

$$\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 + \cdots + \alpha_n\vec{x}_n = \vec{0}$$

is $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$*. Otherwise, we say the vectors are* linearly dependent.

Example 69. *The vectors* $\left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ *and* $\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ *are linearly independent.*

Solution: Let's try:

$$\alpha_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \alpha_2\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ 2\alpha_1 + \alpha_2 \end{bmatrix} = \vec{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

So $\alpha_1 = 0$, and then it is clear that $\alpha_2 = 0$ as well. In other words, the vectors are linearly independent.

If a set of vectors is linearly dependent, that is, we have an expression of the form

$$\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 + \cdots + \alpha_n\vec{x}_n = 0$$

with some of the α_j's are nonzero, then we can solve for one vector in terms of the others. Suppose $\alpha_1 \neq 0$. Since $\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 + \cdots + \alpha_n\vec{x}_n = \vec{0}$, then

$$\vec{x}_1 = \frac{-\alpha_2}{\alpha_1}\vec{x}_2 + \frac{-\alpha_3}{\alpha_1}\vec{x}_3 + \cdots + \frac{-\alpha_n}{\alpha_1}\vec{x}_n.$$

For example,

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

and so

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

In particular, this tells us that if two vectors $\vec{y}_1$ and $\vec{y}_2$ are linearly dependent, then it must be the case that $\vec{y}_1 = A\vec{y}_2$ for some constant A, namely $A = \dfrac{-\alpha_2}{\alpha_1}$ where $\alpha_1\vec{y}_1 + \alpha_2\vec{y}_2 = \vec{0}$. If this is not the case, then the vectors are linearly independent. For more than two vectors, the process is more complicated and involves solving a system of linear equations, which we will deal with in § .

Matrices

The next object we need to define here is a *matrix.*

Definition 13. *In general, an $m \times n$ matrix A is a rectangular array of mn numbers,*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

An $m \times n$ matrix indicates that it will have m rows and n columns.

Matrices, just like vectors, are generally written with square brackets on the outside, although some books will use parentheses for this. The convention for notation is that matrices will be denoted by capital letters (A) and the individual *entries* of the matrix, the numbers that make it up, will be denoted using lowercase letters (a_{ij}) where the first number i indicates which row of the matrix we are talking about, and the second number j indicates which column. For example, in the matrix

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

we could talk about the entire matrix usint A, but would also have that $a_{21} = -2$ and $a_{33} = 5$.

Note that an $m \times 1$ matrix is just a column vector, so in terms of the basic structure, matrices are an extension of vectors. However, they can be used for so much more, as we will see in future sections.

Another way to view matrices is as a set of column vectors all laid out side-by-side. If we have $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$, three different four component vectors, we can form a 4×3 matrix B as

$$B = [\vec{v}_1 \mid \vec{v}_2 \mid \vec{v}_3]$$

that uses each of the given vectors as a column of the matrix. In this case, the vertical lines are used to indicate that this is actually a matrix, because each of the entries given there are vectors, not just individual numbers. If we wanted to write a 1×3 matrix this way, these vertical lines will not be included.

We will go into more properties of matrices and the operations we can perform on them in § . To conclude this section though, we will look at one other way that matrices come about, and that is as the representation of a linear map.

Linear mappings and matrices

A *vector-valued function* F is a rule that takes a vector $\vec{x}$ and returns another vector $\vec{y}$. For example, F could be a scaling that doubles the size of vectors:

$$F(\vec{x}) = 2\vec{x}.$$

For example,

$$F\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right) = 2\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}.$$

If F is a mapping that takes vectors in $\mathbb{R}^2$ to $\mathbb{R}^2$ (such as the above), we write

$$F \colon \mathbb{R}^2 \to \mathbb{R}^2.$$

The words *function* and *mapping* are used rather interchangeably, although more often than not, *mapping* is used when talking about a vector-valued function, and the word *function* is often used when the function is scalar-valued.

A beginning student of mathematics (and many a seasoned mathematician), that sees an expression such as

$$f(3x + 8y)$$

yearns to write

$$3f(x) + 8f(y).$$

After all, who hasn't wanted to write $\sqrt{x+y} = \sqrt{x} + \sqrt{y}$ or something like that at some point in their mathematical lives. Wouldn't life be simple if we could do that? Of course we can't always do that (for example, not with the square roots!) It turns out there are many functions where we can do exactly the above. Such functions are called *linear.*

A mapping $F \colon \mathbb{R}^n \to \mathbb{R}^m$ is called *linear* if
$$F(\vec{x}+\vec{y}) = F(\vec{x}) + F(\vec{y}),$$
for any vectors $\vec{x}$ and $\vec{y}$, and also
$$F(\alpha\vec{x}) = \alpha F(\vec{x}),$$
for any scalar α. The F we defined above that doubles the size of all vectors is linear. Let us check:
$$F(\vec{x}+\vec{y}) = 2(\vec{x}+\vec{y}) = 2\vec{x}+2\vec{y} = F(\vec{x}) + F(\vec{y}),$$
and also
$$F(\alpha\vec{x}) = 2\alpha\vec{x} = \alpha 2\vec{x} = \alpha F(\vec{x}).$$

We also call a linear function a *linear transformation*. If you want to be really fancy and impress your friends, you can call it a *linear operator*.

When a mapping is linear we often do not write the parentheses. We write simply
$$F\vec{x}$$
instead of $F(\vec{x})$. We do this because linearity means that the mapping F behaves like multiplying $\vec{x}$ by "something." That something is a matrix.

Now how does a matrix A relate to a linear mapping? Well a matrix tells you where certain special vectors go. Let's give a name to those certain vectors. The *standard basis vectors* of $\mathbb{R}^n$ are
$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \qquad \cdots, \qquad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

For example, in $\mathbb{R}^3$ these vectors are
$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

You may recall from calculus of several variables that these are sometimes called $\vec{\imath}$, $\vec{\jmath}$, $\vec{k}$.

The reason these are called a *basis* is that every other vector can be written as a *linear combination* of them. For example, in $\mathbb{R}^3$ the vector $(4,5,6)$ can be written as
$$4\vec{e}_1 + 5\vec{e}_2 + 6\vec{e}_3 = 4\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 5\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 6\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

Keep this idea of linear combinations of vectors in mind; we'll see a lot more of it later.

So how does a matrix represent a linear mapping? Well, the columns of the matrix are the vectors where A as a linear mapping takes $\vec{e}_1$, $\vec{e}_2$, etc. For example, consider
$$M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

As a linear mapping $M \colon \mathbb{R}^2 \to \mathbb{R}^2$ takes $\vec{e}_1 = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} 1 \\ 3 \end{smallmatrix}\right]$ and $\vec{e}_2 = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} 2 \\ 4 \end{smallmatrix}\right]$. In other words,
$$M\vec{e}_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \qquad \text{and} \qquad M\vec{e}_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

More generally, if we have an $n \times m$ matrix A, that is we have n rows and m columns, then the mapping $A \colon \mathbb{R}^m \to \mathbb{R}^n$ takes $\vec{e}_j$ to the j^{th} column of A. For example,
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{bmatrix}$$

represents a mapping from $\mathbb{R}^5$ to $\mathbb{R}^3$ that does

$$A\vec{e}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}, \qquad A\vec{e}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix}, \qquad A\vec{e}_3 = \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}, \qquad A\vec{e}_4 = \begin{bmatrix} a_{14} \\ a_{24} \\ a_{34} \end{bmatrix}, \qquad A\vec{e}_5 = \begin{bmatrix} a_{15} \\ a_{25} \\ a_{35} \end{bmatrix}.$$

But what if I have another vector $\vec{x}$? Where does it go? Well we use linearity. First write the vector as a linear combination of the standard basis vectors:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 + x_4\vec{e}_4 + x_5\vec{e}_5.$$

Then

$$A\vec{x} = A(x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 + x_4\vec{e}_4 + x_5\vec{e}_5) = x_1A\vec{e}_1 + x_2A\vec{e}_2 + x_3A\vec{e}_3 + x_4A\vec{e}_4 + x_5A\vec{e}_5.$$

If we know where A takes all the basis vectors, we know where it takes all vectors.

As an example, suppose M is the 2×2 matrix from above, and suppose we wish to find

$$M \begin{bmatrix} -2 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ 0.1 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 0.1 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -1.8 \\ -5.6 \end{bmatrix}.$$

Every linear mapping from $\mathbb{R}^m$ to $\mathbb{R}^n$ can be represented by an $n \times m$ matrix. You just figure out where it takes the standard basis vectors. Conversely, every $n \times m$ matrix represents a linear mapping. Hence, we may think of matrices being linear mappings, and linear mappings being matrices.

Or can we? In this book we study mostly linear differential operators, and linear differential operators are linear mappings, although they are not acting on $\mathbb{R}^n$, but on an infinite-dimensional space of functions:

$$Lf = g$$

for a function f we get a function g, and L is linear in the sense that

$$L(f + h) = Lf + Lh, \qquad \text{and} \qquad L(\alpha f) = \alpha Lf.$$

for any number (scalars) α and all functions f and h.

So the answer is not really. But if we consider vectors in finite-dimensional spaces $\mathbb{R}^n$ then yes, every linear mapping is a matrix. We have mentioned at the beginning of this section, that we can "make everything a vector." That's not strictly true, but it is true approximately. Those "infinite-dimensional" spaces of functions can be approximated by a finite-dimensional space, and then linear operators are just matrices. So approximately, this is true. And as far as actual computations that we can do on a computer, we can work only with finitely many dimensions anyway. If you ask a computer or your calculator to plot a function, it samples the function at finitely many points and then connects the dots [5]. It does not actually give you infinitely many values. So the way that you have been using the computer or your calculator so far has already been a certain approximation of the space of functions by a finite-dimensional space.

[5] In Matlab, you may have noticed that to plot a function, we take a vector of inputs, ask Matlab to compute the corresponding vector of values of the function, and then we ask it to plot the result.

`linearAlgebra/practice/vectorMaps-practice1.tex`

Practice for Forced Oscillations

Why?

Exercise 388 *On a piece of graph paper draw the vectors:*

a) $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$ b) $\begin{bmatrix} -2 \\ -4 \end{bmatrix}$ c) $(3, -4)$

Exercise 389 *On a piece of graph paper draw the vector $(1, 2)$ starting at (based at) the given point:*

a) based at $(0, 0)$ b) based at $(1, 2)$ c) based at $(0, -1)$

Exercise 390 *On a piece of graph paper draw the following operations. Draw and label the vectors involved in the operations as well as the result:*

a) $\begin{bmatrix} 1 \\ -4 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ b) $\begin{bmatrix} -3 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ c) $3 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Exercise 391 *Compute the magnitude of*

a) $\begin{bmatrix} 7 \\ 2 \end{bmatrix}$ b) $\begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$ c) $(1, 3, -4)$

Exercise 392 *Compute the magnitude of*

a) $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ b) $\begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$ c) $(-2, 1, -2)$

Exercise 393 *Compute*

a) $\begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 7 \\ -8 \end{bmatrix}$ b) $\begin{bmatrix} -2 \\ 3 \end{bmatrix} - \begin{bmatrix} 6 \\ -4 \end{bmatrix}$ c) $-\begin{bmatrix} -3 \\ 2 \end{bmatrix}$

d) $4 \begin{bmatrix} -1 \\ 5 \end{bmatrix}$ e) $5 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 9 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ f) $3 \begin{bmatrix} 1 \\ -8 \end{bmatrix} - 2 \begin{bmatrix} 3 \\ -1 \end{bmatrix}$

Exercise 394 *Compute*

Author(s): Matthew Charnley and Jason Nowell

a) $\begin{bmatrix}3\\1\end{bmatrix}+\begin{bmatrix}6\\-3\end{bmatrix}$ b) $\begin{bmatrix}-1\\2\end{bmatrix}-\begin{bmatrix}2\\-1\end{bmatrix}$ c) $-\begin{bmatrix}-5\\3\end{bmatrix}$

d) $2\begin{bmatrix}-2\\4\end{bmatrix}$ e) $3\begin{bmatrix}1\\0\end{bmatrix}+7\begin{bmatrix}0\\1\end{bmatrix}$ f) $2\begin{bmatrix}2\\-3\end{bmatrix}-6\begin{bmatrix}2\\-1\end{bmatrix}$

Exercise 395 *Find the unit vector in the direction of the given vector*

a) $\begin{bmatrix}1\\-3\end{bmatrix}$ b) $\begin{bmatrix}2\\1\\-1\end{bmatrix}$ c) $(3,1,-2)$

Exercise 396 *Find the unit vector in the direction of the given vector*

a) $\begin{bmatrix}-1\\1\end{bmatrix}$ b) $\begin{bmatrix}1\\-1\\2\end{bmatrix}$ c) $(2,-5,2)$

Exercise 397 *If $\vec{x}=(1,2)$ and $\vec{y}$ are added together, we find $\vec{x}+\vec{y}=(0,2)$. What is $\vec{y}$?*

Exercise 398 *If $\vec{v}=(1,-4,3)$ and $\vec{w}=(-2,3,-1)$, compute $3\vec{v}-2\vec{w}$ and $4\vec{w}+\vec{v}$.*

Exercise 399 *Write $(1,2,3)$ as a linear combination of the standard basis vectors $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$.*

Exercise 400 *Determine if the following sets of vectors are linearly independent.*

a) $\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$ b) $\left\{\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}2\\2\\2\end{bmatrix}\right\}$

c) $\left\{\begin{bmatrix}1\\2\end{bmatrix},\begin{bmatrix}-1\\3\end{bmatrix},\begin{bmatrix}1\\1\end{bmatrix}\right\}$ d) $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\0\end{bmatrix}\right\}$

e) $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\-1\end{bmatrix},\begin{bmatrix}0\\0\\2\end{bmatrix}\right\}$ f) $\left\{\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\3\\-1\end{bmatrix},\begin{bmatrix}0\\-2\\1\end{bmatrix}\right\}$

Exercise 401 *If the magnitude of $\vec{x}$ is 4, what is the magnitude of*

a) $0\vec{x}$ b) $3\vec{x}$ c) $-\vec{x}$ d) $-4\vec{x}$ e) $\vec{x}+\vec{x}$ f) $\vec{x}-\vec{x}$

Exercise 402 *If the magnitude of $\vec{x}$ is 5, what is the magnitude of*

a) $4\vec{x}$ b) $-2\vec{x}$ c) $-4\vec{x}$

Exercise 403 *Suppose a linear mapping* $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ *takes* $(1,0)$ *to* $(2,-1)$ *and it takes* $(0,1)$ *to* $(3,3)$. *Where does it take*

a) $(1,1)$ *b)* $(2,0)$ *c)* $(2,-1)$

Exercise 404 *Suppose a linear mapping* $F\colon \mathbb{R}^3 \to \mathbb{R}^2$ *takes* $(1,0,0)$ *to* $(2,1)$ *and it takes* $(0,1,0)$ *to* $(3,4)$ *and it takes* $(0,0,1)$ *to* $(5,6)$. *Write down the matrix representing the mapping* F.

Exercise 405 *Suppose that a mapping* $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ *takes* $(1,0)$ *to* $(1,2)$, $(0,1)$ *to* $(3,4)$, *and it takes* $(1,1)$ *to* $(0,-1)$. *Explain why* F *is not linear.*

Exercise 406 *Suppose a linear mapping* $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ *takes* $(1,0)$ *to* $(1,-1)$ *and it takes* $(0,1)$ *to* $(2,0)$. *Where does it take*

a) $(1,1)$ *b)* $(0,2)$ *c)* $(1,-1)$

Exercise 407 *Let* P *represent the space of quadratic polynomials in* t*: a point* (a_0, a_1, a_2) *in* P *represents the polynomial* $a_0 + a_1 t + a_2 t^2$. *Consider the derivative* $\frac{d}{dt}$ *as a mapping of* P *to* P, *and note that* $\frac{d}{dt}$ *is linear. Write down* $\frac{d}{dt}$ *as a* 3×3 *matrix.*

Matrix algebra

Matrix algebra

One-by-one matrices

Let us motivate what we want to achieve with matrices. What do real-valued linear mappings look like? A linear function of real numbers that you have seen in calculus is of the form

$$f(x) = mx + b.$$

However, the properties of linear mappings discussed in the previous section are that

$$f(x+y) = f(x) + f(y) \qquad f(ax) = af(x).$$

Plugging in the definition from above gives that

$$\begin{gathered} f(x+y) = m(x+y) + b = mx + my + b \\ f(ax) = m(ax) + b = a(mx) + b \end{gathered}$$

and neither of these match up appropriately, since

$$\begin{gathered} f(x) + f(y) = mx + b + my + b = mx + my + 2b \\ af(x) = a(mx+b) = a(mx) + ab \end{gathered}.$$

In order for these to work, we need to have $b = 0$. Therefore, real-valued linear mappings of the real line, linear functions that eat numbers and spit out numbers, are just multiplications by a number.

Consider a mapping defined by multiplying by a number. Let's call this number α. The mapping then takes x to αx. What we can do is to *add* such mappings. If we have another mapping β, then

$$\alpha x + \beta x = (\alpha + \beta)x.$$

We get a new mapping $\alpha + \beta$ that multiplies x by, well, $\alpha + \beta$. If D is a mapping that doubles things, $Dx = 2x$, and T is a mapping that triples, $Tx = 3x$, then $D + T$ is a mapping that multiplies by 5, $(D+T)x = 5x$.

Similarly we can *compose* such mappings, that is, we could apply one and then the other. We take x, we run it through the first mapping α to get α times x, then we run αx through the second mapping β. In other words,

$$\beta(\alpha x) = (\beta\alpha)x.$$

We just multiply those two numbers. Using our doubling and tripling mappings, if we double and then triple, that is $T(Dx)$ then we obtain $3(2x) = 6x$. The composition TD is the mapping that multiplies by 6. For larger matrices, composition also ends up being a kind of multiplication.

Matrix addition and scalar multiplication

The mappings that multiply numbers by numbers are just 1×1 matrices. The number α above could be written as a matrix $[\alpha]$. So perhaps we would want to do the same things to all matrices that we did to those 1×1 matrices at the start of this section above. First, let us add matrices. If we have a matrix A and a matrix B that are of the same size, say $m \times n$, then they are mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$. The mapping $A + B$ should also be a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$, and it should do the following to vectors:

$$(A+B)\vec{x} = A\vec{x} + B\vec{x}.$$

Learning outcomes: Perform addition and multiplication operations on matrices Compute inverses of 2×2 matrices Identify triangular, diagonal, and symmetric matrices.

Author(s): Matthew Charnley and Jason Nowell

It turns out you just add the matrices element-wise: If the ij^{th} entry of A is a_{ij}, and the ij^{th} entry of B is b_{ij}, then the ij^{th} entry of $A+B$ is $a_{ij}+b_{ij}$. If

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix},$$

then

$$A + B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{bmatrix}.$$

Let us illustrate on a more concrete example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} + \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & -1 \end{bmatrix} = \begin{bmatrix} 1+7 & 2+8 \\ 3+9 & 4+10 \\ 5+11 & 6-1 \end{bmatrix} = \begin{bmatrix} 8 & 10 \\ 12 & 14 \\ 16 & 5 \end{bmatrix}.$$

Let's check that this does the right thing to a vector. Let's use some of the vector algebra that we already know, and regroup things:

$$\begin{aligned}
\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} + \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} &= \left(2 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} - \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \right) + \left(2 \begin{bmatrix} 7 \\ 9 \\ 11 \end{bmatrix} - \begin{bmatrix} 8 \\ 10 \\ -1 \end{bmatrix} \right) \\
&= 2 \left(\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + \begin{bmatrix} 7 \\ 9 \\ 11 \end{bmatrix} \right) - \left(\begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} + \begin{bmatrix} 8 \\ 10 \\ -1 \end{bmatrix} \right) \\
&= 2 \begin{bmatrix} 1+7 \\ 3+9 \\ 5+11 \end{bmatrix} - \begin{bmatrix} 2+8 \\ 4+10 \\ 6-1 \end{bmatrix} = 2 \begin{bmatrix} 8 \\ 12 \\ 16 \end{bmatrix} - \begin{bmatrix} 10 \\ 14 \\ 5 \end{bmatrix} \\
&= \begin{bmatrix} 8 & 10 \\ 12 & 14 \\ 16 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \quad \left(= \begin{bmatrix} 2(8)-10 \\ 2(12)-14 \\ 2(16)-5 \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \\ 27 \end{bmatrix} \right).
\end{aligned}$$

If we replaced the numbers by letters that would constitute a proof! You'll notice that we didn't really have to even compute what the result is to convince ourselves that the two expressions were equal.

If the sizes of the matrices do not match, then addition is not defined. If A is 3×2 and B is 2×5, then we cannot add these matrices. We don't know what that could possibly mean.

It is also useful to have a matrix that when added to any other matrix does nothing. This is the zero matrix, the matrix of all zeros:

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

We often denote the zero matrix by 0 without specifying size. We would then just write $A + 0$, where we just assume that 0 is the zero matrix of the same size as A.

There are really two things we can multiply matrices by. We can multiply matrices by scalars or we can multiply by other matrices. Let us first consider multiplication by scalars. For a matrix A and a scalar α we want αA to be the matrix that accomplishes

$$(\alpha A)\vec{x} = \alpha(A\vec{x}).$$

That is just scaling the result by α. If you think about it, scaling every term in A by α accomplishes just that: If

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}, \qquad \text{then} \qquad \alpha A = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \alpha a_{13} \\ \alpha a_{21} & \alpha a_{22} & \alpha a_{23} \end{bmatrix}.$$

For example,

$$2 \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{bmatrix}.$$

Let us list some properties of matrix addition and scalar multiplication. Denote by 0 the zero matrix, by α, β scalars, and by A, B, C matrices. Then:

$$\begin{aligned} A + 0 &= A = 0 + A, \\ A + B &= B + A, \\ (A + B) + C &= A + (B + C), \\ \alpha(A + B) &= \alpha A + \alpha B, \\ (\alpha + \beta)A &= \alpha A + \beta A. \end{aligned}$$

These rules should look very familiar.

Matrix multiplication

As we mentioned above, composition of linear mappings is also a multiplication of matrices. Suppose A is an $m \times n$ matrix, that is, A takes $\mathbb{R}^n$ to $\mathbb{R}^m$, and B is an $n \times p$ matrix, that is, B takes $\mathbb{R}^p$ to $\mathbb{R}^n$. The composition AB should work as follows

$$AB\vec{x} = A(B\vec{x}).$$

First, a vector $\vec{x}$ in $\mathbb{R}^p$ gets taken to the vector $B\vec{x}$ in $\mathbb{R}^n$. Then the mapping A takes it to the vector $A(B\vec{x})$ in $\mathbb{R}^m$. In other words, the composition AB should be an $m \times p$ matrix. In terms of sizes we should have

$$\text{“} \quad [m \times n]\,[n \times p] = [m \times p]. \quad \text{”}$$

Notice how the middle size must match.

OK, now we know what sizes of matrices we should be able to multiply, and what the product should be. Let us see how to actually compute matrix multiplication. We start with the so-called *dot product* (or *inner product*) of two vectors. Usually this is a row vector multiplied with a column vector of the same size. Dot product multiplies each pair of entries from the first and the second vector and sums these products. The result is a single number. For example,

$$\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

And similarly for larger (or smaller) vectors. A dot product is really a product of two matrices: a $1 \times n$ matrix and an $n \times 1$ matrix resulting in a 1×1 matrix, that is, a number.

Armed with the dot product we define the *product of matrices*. First let us denote by $\operatorname{row}_i(A)$ the i^{th} row of A and by $\operatorname{column}_j(A)$ the j^{th} column of A. For an $m \times n$ matrix A and an $n \times p$ matrix B we can compute the product AB. The matrix AB is an $m \times p$ matrix whose ij^{th} entry is the dot product

$$\operatorname{row}_i(A) \cdot \operatorname{column}_j(B).$$

For example, given a 2×3 and a 3×2 matrix we should end up with a 2×2 matrix:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix}, \tag{29}$$

or with some numbers:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ -7 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 \cdot (-1) + 2 \cdot (-7) + 3 \cdot 1 & 1 \cdot 2 + 2 \cdot 0 + 3 \cdot (-1) \\ 4 \cdot (-1) + 5 \cdot (-7) + 6 \cdot 1 & 4 \cdot 2 + 5 \cdot 0 + 6 \cdot (-1) \end{bmatrix} = \begin{bmatrix} -12 & -1 \\ -33 & 2 \end{bmatrix}.$$

A useful consequence of the definition is that the evaluation $A\vec{x}$ for a matrix A and a (column) vector $\vec{x}$ is also matrix multiplication. That is really why we think of vectors as column vectors, or $n \times 1$ matrices. For example,

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

If you look at the last section, that is precisely the last example we gave.

You should stare at the computation of multiplication of matrices AB and the previous definition of $A\vec{y}$ as a mapping for a moment. What we are doing with matrix multiplication is applying the mapping A to the columns of B. This is usually written as follows. Suppose we write the $n \times p$ matrix $B = [\vec{b}_1 \ \vec{b}_2 \ \cdots \ \vec{b}_p]$, where $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_p$ are the columns of B. Then for an $m \times n$ matrix A,

$$AB = A[\vec{b}_1 \ \vec{b}_2 \ \cdots \ \vec{b}_p] = [A\vec{b}_1 \ A\vec{b}_2 \ \cdots \ A\vec{b}_p].$$

The columns of the $m \times p$ matrix AB are the vectors $A\vec{b}_1, A\vec{b}_2, \ldots, A\vec{b}_p$. For example, in (29), the columns of

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

are

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix}.$$

This is a very useful way to understand what matrix multiplication is. It should also make it easier to remember how to perform matrix multiplication.

We can go one step farther with this connection. The idea of a linear combination of vectors was defined in § : for a set of vectors $\vec{y}_1$, $\vec{y}_2$, ..., $\vec{y}_n$, a linear combination of those vectors is a vector of the form $\alpha_1\vec{y}_1 + \alpha_2\vec{y}_2 + \cdots \alpha_n\vec{y}_n$ for some real numbers α_1, ..., α_n. If we write out the product from the end of the last example, we see

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} = \begin{bmatrix} a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix} = b_{12} \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + b_{22} \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} + b_{32} \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix}.$$

Thus, computing the product of a matrix and a vector gives a new vector which is a linear combination of the columns of the matrix, where the coefficients are the entries in the vector. This gives two connections between matrix multiplication and linear independence and linear combinations.

(a) If there is a solution to the vector equation $A\vec{x} = \vec{b}$, this means that the vector $\vec{b}$ can be written as a linear combination of the columns of A.

(b) If there is a non-zero solution to the equation $A\vec{x} = \vec{0}$, then the columns of A are linearly dependent, and the solution $\vec{x}$ gives the coefficients necessary to give a linear combination that equals zero.

We'll see a lot more about this, and how to determine it, in § .

Some rules of matrix algebra

For multiplication we want an analogue of a 1. That is, we desire a matrix that just leaves everything as it found it. This analogue is the so-called *identity matrix*. The identity matrix is a square matrix with 1s on the main diagonal and zeros everywhere else. It is usually denoted by I. For each size we have a different identity matrix and so sometimes we may denote the size as a subscript. For example, the I_3 would be the 3×3 identity matrix

$$I = I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let us see how the matrix works on a smaller example,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} \cdot 1 + a_{12} \cdot 0 & a_{11} \cdot 0 + a_{12} \cdot 1 \\ a_{21} \cdot 1 + a_{22} \cdot 0 & a_{21} \cdot 0 + a_{22} \cdot 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Multiplication by the identity from the left looks similar, and also does not touch anything.

We have the following rules for matrix multiplication. Suppose that A, B, C are matrices of the correct sizes so that the following make sense. Let α denote a scalar (number). Then

$$\begin{aligned} A(BC) &= (AB)C && \text{(associative law)}, \\ A(B+C) &= AB + AC && \text{(distributive law)}, \\ (B+C)A &= BA + CA && \text{(distributive law)}, \\ \alpha(AB) &= (\alpha A)B = A(\alpha B), && \\ IA &= A = AI && \text{(identity)}. \end{aligned}$$

Example 70. *Let us demonstrate a couple of these rules. For example, the associative law:*

$$\underbrace{\begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix}}_{A} \Bigg(\underbrace{\begin{bmatrix} 4 & 4 \\ 1 & -3 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} -1 & 4 \\ 5 & 2 \end{bmatrix}}_{C} \Bigg) = \underbrace{\begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 16 & 24 \\ -16 & -2 \end{bmatrix}}_{BC} = \underbrace{\begin{bmatrix} -96 & -78 \\ 64 & 52 \end{bmatrix}}_{A(BC)},$$

and

$$\Bigg(\underbrace{\begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 4 & 4 \\ 1 & -3 \end{bmatrix}}_{B} \Bigg) \underbrace{\begin{bmatrix} -1 & 4 \\ 5 & 2 \end{bmatrix}}_{C} = \underbrace{\begin{bmatrix} -9 & -21 \\ 6 & 14 \end{bmatrix}}_{AB} \underbrace{\begin{bmatrix} -1 & 4 \\ 5 & 2 \end{bmatrix}}_{C} = \underbrace{\begin{bmatrix} -96 & -78 \\ 64 & 52 \end{bmatrix}}_{(AB)C}.$$

Or how about multiplication by scalars:

$$10 \Bigg(\underbrace{\begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 4 & 4 \\ 1 & -3 \end{bmatrix}}_{B} \Bigg) = 10 \underbrace{\begin{bmatrix} -9 & -21 \\ 6 & 14 \end{bmatrix}}_{AB} = \underbrace{\begin{bmatrix} -90 & -210 \\ 60 & 140 \end{bmatrix}}_{10(AB)},$$

$$\Bigg(10 \underbrace{\begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix}}_{A} \Bigg) \underbrace{\begin{bmatrix} 4 & 4 \\ 1 & -3 \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} -30 & 30 \\ 20 & -20 \end{bmatrix}}_{10A} \underbrace{\begin{bmatrix} 4 & 4 \\ 1 & -3 \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} -90 & -210 \\ 60 & 140 \end{bmatrix}}_{(10A)B},$$

and

$$\underbrace{\begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix}}_{A} \Bigg(10 \underbrace{\begin{bmatrix} 4 & 4 \\ 1 & -3 \end{bmatrix}}_{B} \Bigg) = \underbrace{\begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 40 & 40 \\ 10 & -30 \end{bmatrix}}_{10B} = \underbrace{\begin{bmatrix} -90 & -210 \\ 60 & 140 \end{bmatrix}}_{A(10B)}.$$

A multiplication rule you have used since primary school on numbers is quite conspicuously missing for matrices. That is, matrix multiplication is not commutative. Firstly, just because AB makes sense, it may be that BA is not even defined. For example, if A is 2×3, and B is 3×4, the we can multiply AB but not BA.

Even if AB and BA are both defined, does not mean that they are equal. For example, take $A = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$ and $B = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right]$:

$$AB = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \qquad \neq \qquad \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = BA.$$

Inverse

A couple of other algebra rules you know for numbers do not quite work on matrices:

(a) $AB = AC$ does not necessarily imply $B = C$, even if A is not 0.

(b) $AB = 0$ does not necessarily mean that $A = 0$ or $B = 0$.

For example:

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

To make these rules hold, we do not just need one of the matrices to not be zero, we would need to "divide" by a matrix. This is where the *matrix inverse* comes in.

Definition 14. *Suppose that A and B are $n \times n$ matrices such that*

$$AB = I = BA.$$

Then we call B the inverse of A and we denote B by A^{-1}.

If the inverse of A exists, then we say A is invertible. *If A is not invertible, we say A is* singular.

Perhaps not surprisingly, $(A^{-1})^{-1} = A$, since if the inverse of A is B, then the inverse of B is A.

If $A = [a]$ is a 1×1 matrix, then A^{-1} is $a^{-1} = \dfrac{1}{a}$. That is where the notation comes from. The computation is not nearly as simple when A is larger.

The proper formulation of the cancellation rule is:

If A is invertible, then $AB = AC$ implies $B = C$.

The computation is what you would do in regular algebra with numbers, but you have to be careful never to commute matrices:

$$\begin{aligned} AB &= AC, \\ A^{-1}AB &= A^{-1}AC, \\ IB &= IC, \\ B &= C. \end{aligned}$$

And similarly for cancellation on the right:

If A is invertible, then $BA = CA$ implies $B = C$.

The rule says, among other things, that the inverse of a matrix is unique if it exists: If $AB = I = AC$, then A is invertible and $B = C$.

We will see later how to compute an inverse of a matrix in general. For now, let us note that there is a simple formula for the inverse of a 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

For example:

$$\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}^{-1} = \frac{1}{1 \cdot 4 - 1 \cdot 2} \begin{bmatrix} 4 & -1 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1/2 \\ -1 & 1/2 \end{bmatrix}.$$

Let's try it:

$$\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 2 & -1/2 \\ -1 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 2 & -1/2 \\ -1 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Just as we cannot divide by every number, not every matrix is invertible. In the case of matrices however we may have singular matrices that are not zero. For example,

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

is a singular matrix. But didn't we just give a formula for an inverse? Let us try it:

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}^{-1} = \frac{1}{1 \cdot 2 - 1 \cdot 2} \begin{bmatrix} 2 & -1 \\ -2 & 1 \end{bmatrix} = ?$$

We get into a bit of trouble; we are trying to divide by zero.

So a 2×2 matrix A is invertible whenever

$$ad - bc \neq 0$$

and otherwise it is singular. The expression $ad - bc$ is called the *determinant* and we will look at it more carefully in a later section. There is a similar expression for a square matrix of any size.

Special types of matrices

A simple (and surprisingly useful) type of a square matrix is a so-called *diagonal matrix*. It is a matrix whose entries are all zero except those on the main diagonal from top left to bottom right. For example a 4×4 diagonal matrix is of the form

$$\begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix}.$$

Such matrices have nice properties when we multiply by them. If we multiply them by a vector, they multiply the k^{th} entry by d_k. For example,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \cdot 4 \\ 2 \cdot 5 \\ 3 \cdot 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \\ 18 \end{bmatrix}.$$

Similarly, when they multiply another matrix from the left, they multiply the k^{th} row by d_k. For example,

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

On the other hand, multiplying on the right, they multiply the columns:

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

And it is really easy to multiply two diagonal matrices together:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 & 0 & 0 \\ 0 & 2 \cdot 3 & 0 \\ 0 & 0 & 3 \cdot (-1) \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

For this last reason, they are easy to invert, you simply invert each diagonal element:

$$\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}^{-1} = \begin{bmatrix} d_1^{-1} & 0 & 0 \\ 0 & d_2^{-1} & 0 \\ 0 & 0 & d_3^{-1} \end{bmatrix}.$$

Let us check an example

$$\underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}^{-1}}_{A^{-1}} \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}}_{A^{-1}} \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{I}.$$

It is no wonder that the way we solve many problems in linear algebra (and in differential equations) is to try to reduce the problem to the case of diagonal matrices.

Another type of matrix that has similarly nice properties are *triangular* matrices. A matrix is *upper triangular* if all of the entries below the diagonal are zero. For a 3×3 matrix, an upper triangular matrix looks like

$$\begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \end{bmatrix}$$

where the $*$ can be any number. Similarly, a *lower triangular* matrix is one where all of the entries above the diagonal are zero, or, for a 3×3 matrix, something that looks like

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

A matrix that is both upper and lower triangular is diagonal, because only the entries on the diagonal can be non-zero.

Transpose

Vectors do not always have to be column vectors, that is just a convention. Swapping rows and columns is from time to time needed. The operation that swaps rows and columns is the so-called *transpose*. The transpose of A is denoted by A^T. Example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

So transpose takes an $m \times n$ matrix to an $n \times m$ matrix.

A key fact about the transpose is that if the product AB makes sense then $B^T A^T$ also makes sense, at least from the point of view of sizes. In fact, we get precisely the transpose of AB. That is:

$$(AB)^T = B^T A^T.$$

For example,

$$\left(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 2 & -2 \end{bmatrix} \right)^T = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

It is left to the reader to verify that computing the matrix product on the left and then transposing is the same as computing the matrix product on the right.

If we have a column vector $\vec{x}$ to which we apply a matrix A and we transpose the result, then the row vector $\vec{x}^T$ applies to A^T from the left:

$$(A\vec{x})^T = \vec{x}^T A^T.$$

Another place where transpose is useful is when we wish to apply the dot product[1] to two column vectors:

$$\vec{x} \cdot \vec{y} = \vec{y}^T \vec{x}.$$

That is the way that one often writes the dot product in software.

We say a matrix A is *symmetric* if $A = A^T$. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$$

is a symmetric matrix. Notice that a symmetric matrix is always square, that is, $n \times n$. Symmetric matrices have many nice properties[2], and come up quite often in applications.

To end the section, we notice how $A\vec{x}$ can be written more succinctly. Suppose

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \qquad \text{and} \qquad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Then

$$A\vec{x} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{bmatrix}.$$

For example,

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

In other words, you take a row of the matrix, you multiply them by the entries in your vector, you add things up, and that's the corresponding entry in the resulting vector.

[1] As a side note, mathematicians write $\vec{y}^T \vec{x}$ and physicists write $\vec{x}^T \vec{y}$. Shhh…don't tell anyone, but the physicists are probably right on this.

[2] Although so far we have not learned enough about matrices to really appreciate them.

`linearAlgebra/practice/matrixAlgebra-practice1.tex`

Practice for Matrix Algebra

Why?

Exercise 408 *Add the following matrices*

a) $\begin{bmatrix} -1 & 2 & 2 \\ 5 & 8 & -1 \end{bmatrix} + \begin{bmatrix} 3 & 2 & 3 \\ 8 & 3 & 5 \end{bmatrix}$

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

Exercise 409 *Add the following matrices*

a) $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix} + \begin{bmatrix} 5 & 3 & 4 \\ 1 & 2 & 5 \end{bmatrix}$

b) $\begin{bmatrix} 6 & -2 & 3 \\ 7 & 3 & 3 \\ 8 & -1 & 2 \end{bmatrix} + \begin{bmatrix} -1 & -1 & -3 \\ 6 & 7 & 3 \\ -9 & 4 & -1 \end{bmatrix}$

Exercise 410 *Compute*

a) $3\begin{bmatrix} 0 & 3 \\ -2 & 2 \end{bmatrix} + 6\begin{bmatrix} 1 & 5 \\ -1 & 5 \end{bmatrix}$

b) $2\begin{bmatrix} -3 & 1 \\ 2 & 2 \end{bmatrix} - 3\begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}$

Exercise 411 *Compute*

a) $2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + 3\begin{bmatrix} -1 & 3 \\ 1 & 2 \end{bmatrix}$

b) $3\begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} - 2\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$

Exercise 412 *Multiply the following matrices*

a) $\begin{bmatrix} -1 & 2 \\ 3 & 1 \\ 5 & 8 \end{bmatrix}\begin{bmatrix} 3 & -1 & 3 & 1 \\ 8 & 3 & 2 & -3 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 1 \\ 1 & 0 & 3 \end{bmatrix}\begin{bmatrix} 2 & 3 & 1 & 7 \\ 1 & 2 & 3 & -1 \\ 1 & -1 & 3 & 0 \end{bmatrix}$

c) $\begin{bmatrix} 4 & 1 & 6 & 3 \\ 5 & 6 & 5 & 0 \\ 4 & 6 & 6 & 0 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 1 & 2 \\ 3 & 5 \\ 5 & 6 \end{bmatrix}$

d) $\begin{bmatrix} 1 & 1 & 4 \\ 0 & 5 & 1 \end{bmatrix}\begin{bmatrix} 2 & 2 \\ 1 & 0 \\ 6 & 4 \end{bmatrix}$

Exercise 413 *Multiply the following matrices*

a) $\begin{bmatrix} 2 & 1 & 4 \\ 3 & 4 & 4 \end{bmatrix}\begin{bmatrix} 2 & 4 \\ 6 & 3 \\ 3 & 5 \end{bmatrix}$

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

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

d) $\begin{bmatrix} -2 & -2 \\ 5 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 0 & 3 \\ 1 & 3 \end{bmatrix}$

Author(s): Matthew Charnley and Jason Nowell

Exercise 414

a) *How must the dimensions of two matrices line up in order to multiply them together? If they can be multiplied, what is the dimension of the product?*

b) *If A is a 3×2 matrix and the product AB is a 3×4 matrix, then what are the dimensions of B?*

c) *If A is a 5×3 matrix, is it possible to find a matrix B so that the product AB is a 4×3 matrix? What about a matrix C so that the product CA is a 4×3 matrix?*

Exercise 415 *Assume that A is a 3×4 matrix.*

a) *What must the dimensions of B be in order for the product AB to be defined?*

b) *What must the dimensions of B be in order for the product BA to be defined?*

c) *What about if we want to compute ABA or BAB?*

Exercise 416 *Complete Exercise but with A being a 2×2 matrix.*

Exercise 417 *Compute the inverse of the given matrices*

a) $\begin{bmatrix} -3 \end{bmatrix}$ b) $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 4 \\ 1 & 3 \end{bmatrix}$ d) $\begin{bmatrix} 2 & 2 \\ 1 & 4 \end{bmatrix}$

Exercise 418 *Compute the inverse of the given matrices*

a) $\begin{bmatrix} 2 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ d) $\begin{bmatrix} 4 & 2 \\ 4 & 4 \end{bmatrix}$

Exercise 419 *Compute the inverse of the given matrices*

a) $\begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0.01 & 0 \\ 0 & 0 & 0 & -5 \end{bmatrix}$

Exercise 420 *Compute the inverse of the given matrices*

a) $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ b) $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ c) $\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0.1 \end{bmatrix}$

Exercise 421 *Consider the matrices*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}.$$

a) *Compute the products AB and AC.*

b) *Verify that for these matrices $AB = AC$, but $B \neq C$.*

Exercise 422 *Consider the matrices*

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

Verify that AB and BA are both equal to zero, but neither of the matrices A and B are zero.

Exercise 423

a) *Let A be a 3×4 matrix. What dimension does the vector $\vec{v}$ need to be in order for the product $A\vec{v}$ to be defined? If this product is defined, what is the dimension of the product $A\vec{v}$?*

b) *Let B be a 3×3 matrix. What dimension does the vector $\vec{v}$ need to be in order for the product $B\vec{v}$ to be defined? If this product is defined, what is the dimension of the product $B\vec{v}$?*

Elimination

We discuss Elimination

Linear systems of equations

One application of matrices is to solve systems of linear equations[1]. Consider the following system of linear equations

$$\begin{aligned} 2x_1 + 2x_2 + 2x_3 &= 2, \\ x_1 + x_2 + 3x_3 &= 5, \\ x_1 + 4x_2 + x_3 &= 10. \end{aligned} \tag{30}$$

There is a systematic procedure called *elimination* to solve such a system. In this procedure, we attempt to eliminate each variable from all but one equation. We want to end up with equations such as $x_3 = 2$, where we can just read off the answer.

We write a system of linear equations as a matrix equation:

$$A\vec{x} = \vec{b}.$$

The system (30) is written as

$$\underbrace{\begin{bmatrix} 2 & 2 & 2 \\ 1 & 1 & 3 \\ 1 & 4 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 2 \\ 5 \\ 10 \end{bmatrix}}_{\vec{b}}.$$

If we knew the inverse of A, then we would be done; we would simply solve the equation:

$$\vec{x} = A^{-1}A\vec{x} = A^{-1}\vec{b}.$$

Well, but that is part of the problem, we do not know how to compute the inverse for matrices bigger than 2×2. We will see later that to compute the inverse we are really solving $A\vec{x} = \vec{b}$ for several different $\vec{b}$. In other words, we will need to do elimination to find A^{-1}. In addition, we may wish to solve $A\vec{x} = \vec{b}$ even if A is not invertible, or perhaps not even square.

Let us return to the equations themselves and see how we can manipulate them. There are a few operations we can perform on the equations that do not change the solution. First, perhaps an operation that may seem stupid, we can swap two equations in (30):

$$\begin{aligned} x_1 + x_2 + 3x_3 &= 5, \\ 2x_1 + 2x_2 + 2x_3 &= 2, \\ x_1 + 4x_2 + x_3 &= 10. \end{aligned}$$

Clearly these new equations have the same solutions x_1, x_2, x_3. A second operation is that we can multiply an equation by a nonzero number. For example, we multiply the third equation in (30) by 3:

$$\begin{aligned} 2x_1 + 2x_2 + 2x_3 &= 2, \\ x_1 + x_2 + 3x_3 &= 5, \\ 3x_1 + 12x_2 + 3x_3 &= 30. \end{aligned}$$

Finally we can add a multiple of one equation to another equation. For example, we add 3 times the third equation in (30) to the second equation:

$$\begin{aligned} 2x_1 + \qquad 2x_2 + \qquad 2x_3 &= 2, \\ (1+3)x_1 + (1+12)x_2 + (3+3)x_3 &= 5 + 30, \\ x_1 + \qquad 4x_2 + \qquad x_3 &= 10. \end{aligned}$$

Learning outcomes: Write a system of linear equations in matrix form Use row reduction to put a matrix into row echlon form or reduced row echelon form Determine whether a system of linear equations has no solution, one solution, or infinitely many solutions.

Author(s): Matthew Charnley and Jason Nowell

[1] Although perhaps we have this backwards, quite often we solve a linear system of equations to find out something about matrices, rather than vice versa.

The same x_1, x_2, x_3 should still be solutions to the new equations. These were just examples; we did not get any closer to the solution. We must to do these three operations in some more logical manner, but it turns out these three operations suffice to solve every linear equation.

The first thing is to write the equations in a more compact manner. Given

$$A\vec{x} = \vec{b},$$

we write down the so-called *augmented matrix*

$$[A \mid \vec{b}],$$

where the vertical line is just a marker for us to know where the "right-hand side" of the equation starts. For example, for the system (30) the augmented matrix is

$$\left[\begin{array}{ccc|c} 2 & 2 & 2 & 2 \\ 1 & 1 & 3 & 5 \\ 1 & 4 & 1 & 10 \end{array}\right].$$

The entire process of elimination, which we will describe, is often applied to any sort of matrix, not just an augmented matrix. Simply think of the matrix as the 3×4 matrix

$$\begin{bmatrix} 2 & 2 & 2 & 2 \\ 1 & 1 & 3 & 5 \\ 1 & 4 & 1 & 10 \end{bmatrix}.$$

Row echelon form and elementary operations

We apply the three operations above to the matrix. We call these the *elementary operations* or *elementary row operations*.

Definition 15. *The elementary row operations on a matrix are:*

(a) *Swap two rows.*

(b) *Multiply a row by a nonzero number.*

(c) *Add a multiple of one row to another row.*

Note that these are the same three operations that we could do with equations to try to solve them earliner in this section. We run these operations until we get into a state where it is easy to read off the answer, or until we get into a contradiction indicating no solution.

More specifically, we run the operations until we obtain the so-called *row echelon form*. Let us call the first (from the left) nonzero entry in each row the *leading entry*. A matrix is in *row echelon form* if the following conditions are satisfied:

(a) The leading entry in any row is strictly to the right of the leading entry of the row above.

(b) Any zero rows are below all the nonzero rows.

(c) All leading entries are 1.[2]

A matrix is in *reduced row echelon form* if furthermore the following condition is satisfied.

(a) All the entries above a leading entry are zero.

Example 71. *The following matrices are in row echelon form. The leading entries are marked:*

$$\begin{bmatrix} 1 & 2 & 9 & 3 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & -1 & -3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & -5 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Note that the definition applies to matrices of any size. None of the matrices above are in reduced *row echelon form. For example, in the first matrix none of the entries above the second and third leading entries are zero; they are 9, 3, and 5.*

[2] Some books do not require this and allow the entries to be any non-zero number, putting this as a requirement for reduced row echelon form. I like leaving this here because it makes the process of row reduction seem more defined and algorithmic. They are equivalent though and with or without the 1 requirement can be used to answer the same questions.

The following matrices are in reduced row echelon form. The leading entries are marked:

$$\begin{bmatrix} 1 & 3 & 0 & 8 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The procedure we will describe to find a reduced row echelon form of a matrix is called *Gauss–Jordan elimination*. The first part of it, which obtains a row echelon form, is called *Gaussian elimination* or *row reduction*. For some problems, a row echelon form is sufficient, and it is a bit less work to only do this first part.

To attain the row echelon form we work *systematically*. We go column by column, starting at the first column. We find topmost entry in the first column that is not zero, and we call it the *pivot*. If there is no nonzero entry we move to the next column. We swap rows to put the row with the pivot as the first row. We divide the first row by the pivot to make the pivot entry be a 1. Now look at all the rows below and subtract the correct multiple of the pivot row so that all the entries below the pivot become zero.

After this procedure we forget that we had a first row (it is now fixed), and we forget about the column with the pivot and all the preceding zero columns. Below the pivot row, all the entries in these columns are just zero. Then we focus on the smaller matrix and we repeat the steps above.

It is best shown by example, so let us go back to the example from the beginning of the section. We keep the vertical line in the matrix, even though the procedure works on any matrix, not just an augmented matrix. We start with the first column and we locate the pivot, in this case the first entry of the first column.

$$\left[\begin{array}{ccc|c} 2 & 2 & 2 & 2 \\ 1 & 1 & 3 & 5 \\ 1 & 4 & 1 & 10 \end{array}\right]$$

We multiply the first row by $1/2$.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & 3 & 5 \\ 1 & 4 & 1 & 10 \end{array}\right]$$

We subtract the first row from the second and third row (two elementary operations).

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 4 \\ 0 & 3 & 0 & 9 \end{array}\right]$$

We are done with the first column and the first row for now. We almost pretend the matrix doesn't have the first column and the first row.

$$\left[\begin{array}{ccc|c} * & * & * & * \\ * & 0 & 2 & 4 \\ * & 3 & 0 & 9 \end{array}\right]$$

OK, look at the second column, and notice that now the pivot is in the third row.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 4 \\ 0 & 3 & 0 & 9 \end{array}\right]$$

We swap rows.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 3 & 0 & 9 \\ 0 & 0 & 2 & 4 \end{array}\right]$$

And we divide the pivot row by 3.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 2 & 4 \end{array}\right]$$

We do not need to subtract anything as everything below the pivot is already zero. We move on, we again start ignoring the second row and second column and focus on

$$\left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ * & * & 2 & 4 \end{array}\right].$$

We find the pivot, then divide that row by 2:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 2 & 4 \end{array}\right] \quad \to \quad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

The matrix is now in row echelon form.

The equation corresponding to the last row is $x_3 = 2$. We know x_3 and we could substitute it into the first two equations to get equations for x_1 and x_2. Then we could do the same thing with x_2, until we solve for all 3 variables. This procedure is called *backsubstitution* and we can achieve it via elementary operations. We start from the lowest pivot (leading entry in the row echelon form) and subtract the right multiple from the row above to make all the entries above this pivot zero. Then we move to the next pivot and so on. After we are done, we will have a matrix in reduced row echelon form.

We continue our example. Subtract the last row from the first to get

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

The entries above the pivot in the third row is already zero. So we move onto the next pivot, the one in the second row. We subtract this row from the top row to get

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

The matrix is in reduced row echelon form.

If we now write down the equations for x_1, x_2, x_3, we find

$$x_1 = -4, \qquad x_2 = 3, \qquad x_3 = 2.$$

In other words, we have solved the system.

Example 72. *Solve the following system of equations using row reduction:*

$$\begin{aligned} -x_1 + x_2 + 3x_3 &= 7 \\ -3x_1 \qquad + \ x_3 &= -5 \\ -2x_1 - x_2 &= -4 \end{aligned}$$

Solution: In order to solve this problem, we need to set up the augmented matrix for this system, which is

$$\left[\begin{array}{ccc|c} -1 & 1 & 3 & 7 \\ -3 & 0 & 1 & -5 \\ -2 & -1 & 0 & -4 \end{array}\right]$$

To carry out the process, we need to get a 1 in the top left corner, then work from there. We multiply the first row by -1 to get

$$\left[\begin{array}{ccc|c} 1 & -1 & -3 & -7 \\ -3 & 0 & 1 & -5 \\ -2 & -1 & 0 & -4 \end{array}\right].$$

Next, we want to use row 1 to cancel out the -3 and -2 in column 1. To do this, we add three copies of row 1 to row 2, and two copies of row 1 to row 3 to get the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & -1 & -3 & -7 \\ 0 & -3 & -8 & -26 \\ 0 & -3 & -6 & -18 \end{array}\right].$$

Normally, the next step would be to divide the second row by -3 in order to put a 1 in that pivot spot. However, since both the second and third rows have a -3 in the second column, we can combine these two rows directly without dividing by -3 first. We subtract row 2 from row 3 to get

$$\left[\begin{array}{ccc|c} 1 & -1 & -3 & -7 \\ 0 & -3 & -8 & -26 \\ 0 & 0 & 2 & 8 \end{array}\right]$$

and we can now use this to solve the system. The bottom row says that $2x_3 = 8$, so that $x_3 = 4$. The second row says that $-3x_2 - 8x_3 = -26$, since $x_3 = 4$, we have that $-3x_2 = -26 + 32 = 6$, so $x_2 = -2$. Finally, the first row of the augmented matrix says that $x_1 - x_2 - 3x_3 = -7$. Plugging in our values for x_2 and x_3 gives $x_1 = -7 - 2 + 12 = 3$. Therefore, the solution is

$$x_1 = 3 \quad x_2 = -2 \quad x_3 = 4.$$

Non-unique solutions and inconsistent systems

It is possible that the solution of a linear system of equations is not unique, or that no solution exists. Suppose for a moment that the row echelon form we found was

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Then the last row gives the equation $0x_1 + 0x_2 + 0x_3 = 1$, or $0 = 1$. That is impossible and the equations are *inconsistent*. There is no solution to $A\vec{x} = \vec{b}$.

On the other hand, if we find a row echelon form of

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

then there is no issue with finding solutions. In fact, we will find way too many. Let us continue with backsubstitution (subtracting 3 times the second row from the first) to find the reduced row echelon form and let's mark the pivots.

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

The last row is all zeros; it just says $0 = 0$ and we ignore it. The two remaining equations are

$$x_1 + 2x_2 = -5, \qquad x_3 = 3.$$

Let us solve for the variables that corresponded to the pivots, that is x_1 and x_3 as there was a pivot in the first column and in the third column:

$$\begin{aligned} x_1 &= -2x_2 - 5, \\ x_3 &= 3. \end{aligned}$$

The variable x_2 can be anything you wish and we still get a solution. The x_2 is called a *free variable*. There are infinitely many solutions, one for every choice of x_2. For example, if we pick $x_2 = 0$, then $x_1 = -5$, and $x_3 = 3$ give a solution. But we also get a solution by picking say $x_2 = 1$, in which case $x_1 = -7$ and $x_3 = 3$, or by picking $x_2 = -5$ in which case $x_1 = 5$ and $x_3 = 3$.

The general idea is that if any row has all zeros in the columns corresponding to the variables, but a nonzero entry in the column corresponding to the right-hand side $\vec{b}$, then the system is inconsistent and has no solutions. In other words, the system is inconsistent if you find a pivot on the right side of the vertical line drawn in the augmented matrix. Otherwise, the system is consistent, and at least one solution exists.

If the system is consistent:

(a) If every column corresponding to a variable has a pivot element, then the solution is unique.

(b) If there are columns corresponding to variables with no pivot, then those are *free variables* that can be chosen arbitrarily, and there are infinitely many solutions.

Another way to interpret this idea of free variables is that at the beginning, before you look at the system of equations, all of the variables can be anything, and there are no constraints on them. The equations then give us constraints on these variables, because they give us rules that the variables must satisfy. When we have a row of the augmented matrix that becomes all zeros, it means that the equation that was there is redundant and doesn't add any constraints to the equations. This may result in an *underdetermined* system, which will likely have free variables.

Example 73. *Solve the following two systems of equations, or determine that no solution exists, using row reduction:*

$$\begin{aligned} x_1 - x_2 - 3x_3 &= -3 \\ -x_1 - 2x_2 + 4x_3 &= 6 \\ x_1 + 5x_2 - 5x_3 &= -9 \end{aligned}$$

$$\begin{aligned} x_1 - x_2 - 3x_3 &= -3 \\ -x_1 - 2x_2 + 4x_3 &= 6 \\ x_1 + 5x_2 - 5x_3 &= 1 \end{aligned}$$

Solution: For the first of these systems, we will set up the augmented matrix and proceed through the process like normal. The augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & -1 & -3 & -3 \\ -1 & -2 & 4 & 6 \\ 1 & 5 & -5 & -9 \end{array}\right].$$

Since we already have a 1 in the top-left corner of this matrix, we can use it to cancel the entries in the rest of column 1. We add one copy of row 1 to row 2, and subtract row 1 from row 3 to get the next augmented form matrix as

$$\left[\begin{array}{ccc|c} 1 & -1 & -3 & -3 \\ 0 & -3 & 1 & 3 \\ 0 & 6 & -2 & -6 \end{array}\right].$$

Looking at the matrix here, we see that row 3 is -2 times row 2. Therefore, if we add two copies of row 2 to row 3, we get the augmented matrix

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

Therefore, we have a situation where there are only two pivot columns, and the last row is all zeros. Since there are three variables and column 3 is not a pivot column, we can take x_3 as a free variable. If we do that, the second equation tells us that $-3x_2 + x_3 = 3$, or, since we are taking x_3 as a free variable, we can write $x_2 = -1 + {}^1\!/_3 x_3$. We can then take the first equation, which says that $x_1 - x_2 - 3x_3 = -3$ or, by rearranging

$$x_1 = -3 + x_2 + 3x_3 = -3 + \left(-1 + \frac{1}{3}x_3\right) + 3x_3 = -4 + \frac{10}{3}x_3.$$

This means that for any value of t, our solution is determined by

$$\begin{aligned} x_1 &= -4 + \frac{10}{3}t \\ x_2 &= -1 + \frac{1}{3}t \\ x_3 &= t \end{aligned}.$$

The use of t here is just to separate it from the variable x_3. For example, we could pick $t = 3$, in which case we would get $x_1 = 6$, $x_2 = 0$, $x_3 = 3$.

For the second version of the problem, we again set up the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & -1 & -3 & -3 \\ -1 & -2 & 4 & 6 \\ 1 & 5 & -5 & 1 \end{array}\right].$$

Since the left side matrix part is the same as the previous version, the process of row reducing the matrix is identical to what was done previously. When we carry out this process we get the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & -1 & -3 & -3 \\ 0 & -3 & 1 & 3 \\ 0 & 0 & 0 & 10 \end{array}\right].$$

In this case, we see that the last row corresponds to the equation $0 = 10$ so these equations are inconsistent and do not have a solution.

The point of the above example is to illustrate the fact that whether or not a system is inconsistent or has free variables in the solution depends on the right-hand side of the equation, even if the left-hand side has the same coefficients. We'll see more about why this is in § .

When $\vec{b} = \vec{0}$, we have a so-called *homogeneous matrix equation*

$$A\vec{x} = \vec{0}.$$

There is no need to write an augmented matrix in this case. As the elementary operations do not do anything to a zero column, it always stays a zero column. Moreover, $A\vec{x} = \vec{0}$ always has at least one solution, namely $\vec{x} = \vec{0}$. Such a system is always consistent. It may have other solutions: If you find any free variables, then you get infinitely many solutions. As mentioned in the last section, this is directly connected to linear independence of the columns of A. If there are other solutions, then there are other linear combinations that give $\vec{0}$, and so the columns of A are linearly dependent. Otherwise, if there are no such solutions outside of $\vec{x} = \vec{0}$, then the columns of A are linearly independent.

How would we determine this fact? We don't need to include the zero column on the far right, but we can apply the same operations to the matrix A alone. In this case, we either get a pivot column in every column for the row echelon form, in which case the only solution is $\vec{x} = \vec{0}$, or we get at least one non-pivot column, which means that there are free variables, implying that non-zero solutions exist. This leads to a first equivalence statement we can make about the solution to homogeneous matrix equations.

Theorem 13. *Let A be a matrix. The following statements are equivalent (meaning if any one of them is true, so are all of the other ones):*

(a) *The only solution to the matrix equation $A\vec{x} = \vec{0}$ is $\vec{x} = \vec{0}$.*

(b) *The row echelon form of A has a pivot element in every column.*

(c) *The reduced row echelon form of A is an identity matrix, potentially with rows of zero on the bottom.*

(d) *The columns of A are linearly independent.*

The set of solutions of $A\vec{x} = \vec{0}$ comes up quite often so people give it a name. It is called the *nullspace* or the *kernel* of A. One place where the kernel comes up is invertibility of a square matrix A. If the kernel of A contains a nonzero vector, then it contains infinitely many vectors (there was a free variable). But then it is impossible to invert $\vec{0}$, since infinitely many vectors go to $\vec{0}$, so there is no unique vector that A takes to $\vec{0}$. So if the kernel is nontrivial, that is, if there are any nonzero vectors, in other words, if there are any free variables, or in yet other words, if the row echelon form of A has columns without pivots, then A is not invertible. We will return to this idea later.

`linearAlgebra/practice/elimination-practice1.tex`

Practice for Elimination Method

Why?

Exercise 424 *Compute the reduced row echelon form for the following matrices:*

a) $\begin{bmatrix} 1 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 3 & 3 \\ 6 & -3 \end{bmatrix}$ c) $\begin{bmatrix} 3 & 6 \\ -2 & -3 \end{bmatrix}$ d) $\begin{bmatrix} 6 & 6 & 7 & 7 \\ 1 & 1 & 0 & 1 \end{bmatrix}$

e) $\begin{bmatrix} 9 & 3 & 0 & 2 \\ 8 & 6 & 3 & 6 \\ 7 & 9 & 7 & 9 \end{bmatrix}$ f) $\begin{bmatrix} 2 & 1 & 3 & -3 \\ 6 & 0 & 0 & -1 \\ -2 & 4 & 4 & 3 \end{bmatrix}$ g) $\begin{bmatrix} 6 & 6 & 5 \\ 0 & -2 & 2 \\ 6 & 5 & 6 \end{bmatrix}$ h) $\begin{bmatrix} 0 & 2 & 0 & -1 \\ 6 & 6 & -3 & 3 \\ 6 & 2 & -3 & 5 \end{bmatrix}$

Exercise 425 *Compute the reduced row echelon form for the following matrices:*

a) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix}$ d) $\begin{bmatrix} 1 & -3 & 1 \\ 4 & 6 & -2 \\ -2 & 6 & -2 \end{bmatrix}$

e) $\begin{bmatrix} 2 & 2 & 5 & 2 \\ 1 & -2 & 4 & -1 \\ 0 & 3 & 1 & -2 \end{bmatrix}$ f) $\begin{bmatrix} -2 & 6 & 4 & 3 \\ 6 & 0 & -3 & 0 \\ 4 & 2 & -1 & 1 \end{bmatrix}$ g) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ h) $\begin{bmatrix} 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 5 \end{bmatrix}$

Exercise 426 *Solve (find all solutions), or show no solution exists*

a) $\begin{aligned} 4x_1 + 3x_2 &= -2 \\ -x_1 + x_2 &= 4 \end{aligned}$

b) $\begin{aligned} x_1 + 5x_2 + 3x_3 &= 7 \\ 8x_1 + 7x_2 + 8x_3 &= 8 \\ 4x_1 + 8x_2 + 6x_3 &= 4 \end{aligned}$

c) $\begin{aligned} 4x_1 + 8x_2 + 2x_3 &= 3 \\ -x_1 - 2x_2 + 3x_3 &= 1 \\ 4x_1 + 8x_2 &= 2 \end{aligned}$

d) $\begin{aligned} x + 2y + 3z &= 4 \\ 2x - y + 3z &= 1 \\ 3x + y + 6z &= 6 \end{aligned}$

Exercise 427 *Solve (find all solutions), or show no solution exists*

a) $\begin{aligned} 4x_1 + 3x_2 &= -1 \\ 5x_1 + 6x_2 &= 4 \end{aligned}$

b) $\begin{aligned} 5x + 6y + 5z &= 7 \\ 6x + 8y + 6z &= -1 \\ 5x + 2y + 5z &= 2 \end{aligned}$

c) $\begin{aligned} a + b + c &= -1 \\ a + 5b + 6c &= -1 \\ -2a + 5b + 6c &= 8 \end{aligned}$

d) $\begin{aligned} -2x_1 + 2x_2 + 8x_3 &= 6 \\ x_2 + x_3 &= 2 \\ x_1 + 4x_2 + x_3 &= 7 \end{aligned}$

Author(s): Matthew Charnley and Jason Nowell

Exercise 428 *Solve the system of equations*

$$\begin{aligned} -4x_2 + x_3 + 2x_4 &= 16 \\ 2x_1 + 2x_2 - 4x_3 - 3x_4 &= 1 \\ x_1 + x_2 + 2x_3 + 3x_4 &= 6 \\ 2x_1 - 2x_3 + 4x_4 &= 24 \end{aligned}$$

or determine that no solution exists.

Exercise 429 *Solve the system of equations*

$$\begin{aligned} 3x_2 + 3x_3 + 2x_4 &= 4 \\ 4x_1 + 4x_2 + 2x_3 - 4x_4 &= -26 \\ x_1 - 3x_2 - 2x_3 + 2x_4 &= 1 \\ 3x_1 + 3x_2 + 3x_3 - x_4 &= -14 \end{aligned}$$

or determine that no solution exists.

Exercise 430 *Solve the system of equations*

$$\begin{aligned} 2x_1 + x_2 - x_3 + 4x_4 &= 11 \\ x_1 + 4x_2 - 4x_3 - x_4 &= -7 \\ -2x_1 - 3x_2 + 2x_3 + x_4 &= 11 \\ 3x_1 + x_3 + 4x_4 &= 3 \end{aligned}$$

or determine that no solution exists.

Exercise 431 *Solve the system of equations*

$$\begin{aligned} x_1 - x_3 - 4x_4 &= -3 \\ x_1 + x_2 + x_4 &= 0 \\ x_1 + 3x_2 + 3x_3 - 4x_4 &= -28 \\ 6x_1 + 3x_2 - 4x_3 + 6x_4 &= 25 \end{aligned}$$

or determine that no solution exists.

Exercise 432 *Assume that you are solving a three component linear system of equations via row reduction of an augmented matrix and reach the matrix*

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

What does this mean about the solution to this system of equations?

Exercise 433 *Assume that you are solving a three component linear system of equations via row reduction of an augmented matrix and reach the matrix*

$$\left[\begin{array}{ccc|c} 1 & 1 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

What does this mean about the solution to this system of equations?

Exercise 434 *Assume that you are solving a four component linear system of equations via row reduction of an augmented matrix and reach the matrix*

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 5 & 1 \\ 0 & 2 & 1 & 4 & 2 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 3 & 2 & -1 & 1 \end{array}\right].$$

What is the next step in reducing this matrix? Carry out the rest of this problem to solve the corresponding system of equations.

Exercise 435 *Assume that someone else has provided you the solution to an augmented matrix reduction for solving a system of equations given below*

$$\begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & 3 & 5 \\ 1 & 2 & 4 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & 3 & 5 \\ 0 & 1 & 2 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -9 \end{bmatrix}.$$

Is this work correct? If so, what does this say about the solution(s) to the system? If not, correct the work to solve the system.

Exercise 436 *Find the row echelon form of the matrix A given below. What does this tell you about the solutions to the equation $A\vec{x} = \vec{0}$?*

$$A = \begin{bmatrix} 1 & 2 & 4 \\ -2 & -3 & -6 \\ 1 & 1 & 2 \end{bmatrix}$$

Exercise 437 *Find the row echelon form of the matrix A given below. What does this tell you about the solutions to the equation $A\vec{x} = \vec{0}$?*

$$A = \begin{bmatrix} 1 & -1 & 4 & 1 \\ 2 & -3 & 10 & 2 \\ -3 & 4 & -15 & -4 \\ 3 & -5 & 17 & 3 \end{bmatrix}$$

Linear independence, rank, and dimension

Linear independence, rank, and dimension

Linear independence and rank

As we saw in the § , it is possible to have a set of equations that is redundant; that is, at least one of the equations does not give us any more information or constraints on the variables. In a lot of cases, this either led to inconsistent systems or free variables. We would like to have a better way to talk about this idea, both in terms of systems of equations and matrices in general. The concept we want is that of linear independence. The same concept is useful for differential equations, for example in Chapter **??**.

Definition 16. *Given row or column vectors $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n$, a* linear combination *is an expression of the form*

$$\alpha_1 \vec{y}_1 + \alpha_2 \vec{y}_2 + \cdots + \alpha_n \vec{y}_n,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are all scalars.

For example, $3\vec{y}_1 + \vec{y}_2 - 5\vec{y}_3$ is a linear combination of $\vec{y}_1$, $\vec{y}_2$, and $\vec{y}_3$.

We have seen linear combinations before. The expression

$$A\vec{x}$$

is a linear combination of the columns of A, while

$$\vec{x}^T A = (A^T \vec{x})^T$$

is a linear combination of the rows of A.

The way linear combinations come up in our study of differential equations is similar to the following computation. Suppose that $\vec{x}_1$, $\vec{x}_2$, $\ldots$, $\vec{x}_n$ are solutions to $A\vec{x}_1 = \vec{0}$, $A\vec{x}_2 = \vec{0}$, $\ldots$, $A\vec{x}_n = \vec{0}$. Then the linear combination

$$\vec{y} = \alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \cdots + \alpha_n \vec{x}_n$$

is a solution to $A\vec{y} = \vec{0}$:

$$A\vec{y} = A(\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \cdots + \alpha_n \vec{x}_n) =$$
$$= \alpha_1 A\vec{x}_1 + \alpha_2 A\vec{x}_2 + \cdots + \alpha_n A\vec{x}_n = \alpha_1 \vec{0} + \alpha_2 \vec{0} + \cdots + \alpha_n \vec{0} = \vec{0}.$$

We have seen this computation before in the sense of solutions to homogeneous second order equations. We used $L[y]$ to represent a second order linear differential equation, and showed that if we knew that functions y_1 and y_2 solved

$$L[y_1] = 0 \qquad L[y_2] = 0$$

then $L[c_1 y_1 + c_2 y_2] = 0$ for any constants c_1 and c_2. We did this by showing that

$$L[c_1 y_1 + c_2 y_2] = c_1 L[y_1] + C_2 L[y_2]$$

which mirrors the expression computed above.

Our original question was about when equations are redundant. That is answered by the following definition.

Definition 17. *We say the vectors $\vec{y}_1$, $\vec{y}_2$, $\ldots$, $\vec{y}_n$ are* linearly independent *if the only way to pick α_1, $\alpha_2, \ldots, \alpha_n$ to satisfy*

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \cdots + \alpha_n \vec{x}_n = \vec{0}$$

is $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Otherwise, we say the vectors are linearly dependent.

Learning outcomes: Determine if a set of vectors is linearly independent Compute the rank of a matrix Find a maximal linearly independent subset of a set of vectors Compute a basis of a subspace and the dimension of that subspace.

Author(s): Matthew Charnley and Jason Nowell

If the equations (or their coefficients, as we will see later) are linearly dependent, then they are redundant equations, and not all of them are necessary to define the same solution to the equation. If they are linearly independent, then they are all required.

Example 74. *The vectors* $\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right]$ *and* $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ *are linearly independent.*

Solution: Let's try:

$$\alpha_1 \begin{bmatrix}1\\2\end{bmatrix} + \alpha_2 \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}\alpha_1\\2\alpha_1+\alpha_2\end{bmatrix} = \vec{0} = \begin{bmatrix}0\\0\end{bmatrix}.$$

So $\alpha_1 = 0$, and then it is clear that $\alpha_2 = 0$ as well. In other words, the vectors are linearly independent.

If a set of vectors is linearly dependent, that is, we have an expression of the form

$$\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 + \cdots + \alpha_n\vec{x}_n = 0$$

with some of the α_j's are nonzero, then we can solve for one vector in terms of the others. Suppose $\alpha_1 \neq 0$. Since $\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 + \cdots + \alpha_n\vec{x}_n = \vec{0}$, then

$$\vec{x}_1 = \frac{-\alpha_2}{\alpha_1}\vec{x}_2 - \frac{-\alpha_3}{\alpha_1}\vec{x}_3 + \cdots + \frac{-\alpha_n}{\alpha_1}\vec{x}_n.$$

For example,

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

and so

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

You may have noticed that solving for those α_j's is just solving linear equations, and so you may not be surprised that to check if a set of vectors is linearly independent we use row reduction.

Given a set of vectors, we may not be interested in just finding if they are linearly independent or not, we may be interested in finding a linearly independent subset. Or perhaps we may want to find some other vectors that give the same linear combinations and are linearly independent. The way to figure this out is to form a matrix out of our vectors. If we have row vectors we consider them as rows of a matrix. If we have column vectors we consider them columns of a matrix.

Definition 18. *Given a matrix A, the maximal number of linearly independent rows is called the* rank *of A, and we write "$\operatorname{rank} A$" for the rank.*

For example,

$$\operatorname{rank}\begin{bmatrix}1&1&1\\2&2&2\\-1&-1&-1\end{bmatrix} = 1.$$

The second and third row are multiples of the first one. We cannot choose more than one row and still have a linearly independent set. But what is

$$\operatorname{rank}\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix} \quad = \quad ?$$

That seems to be a tougher question to answer. The first two rows are linearly independent, so the rank is at least two. If we would set up the equations for the α_1, α_2, and α_3, we would find a system with infinitely many solutions. One solution is

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

So the set of all three rows is linearly dependent, the rank cannot be 3. Therefore the rank is 2.

But how can we do this in a more systematic way? We find the row echelon form!

$$\text{Row echelon form of} \quad \begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix} \quad \text{is} \quad \begin{bmatrix}1&2&3\\0&1&2\\0&0&0\end{bmatrix}.$$

The elementary row operations do not change the set of linear combinations of the rows (that was one of the main reasons for defining them as they were). In other words, the span of the rows of the A is the same as the span of the rows of the row echelon form of A. In particular, the number of linearly independent rows is the same. And in the row echelon form, all nonzero rows are linearly independent. This is not hard to see. Consider the two nonzero rows in the example above. Suppose we tried to solve for the α_1 and α_2 in

$$\alpha_1 \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}.$$

Since the first column of the row echelon matrix has zeros except in the first row means that $\alpha_1 = 0$. For the same reason, α_2 is zero. We only have two nonzero rows, and they are linearly independent, so the rank of the matrix is 2. This also tells us that if we were trying to solve the system of equations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= a \\ 4x_1 + 5x_2 + 6x_3 &= b \\ 7x_1 + 8x_2 + 9x_3 &= c \end{aligned}$$

we would get that one row of the reduced augmented matrix has all zeros on the left side, and so this system either has a free variable or is inconsistent, because only two equations here are relevant. We will see more examples of the rank of a matrix once we have more terminology to talk about it.

Subspaces and span

Now, let's consider a different scenario. Assume that we find two vectors that solve $A\vec{x} = 0$. What other vectors also solve this equation? In our discussion of linear combinations, we saw that if $\vec{x}_1$ and $\vec{x}_2$ solve $A\vec{x} = 0$, then so does $A(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2)$ for any constants α_1 and α_2. Thus, all linear combinations will also solve the equation. This leads to the definition of the span of a set of vectors.

Definition 19. *The set of all linear combinations of a set of vectors is called their* span.

$$\operatorname{span}\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\} = \{\textit{Set of all linear combinations of } \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}.$$

Thus, if two vectors solve a homogeneous equation, so does everything in the span of those two vectors. The span of a collection of vectors is an example of a subspace, which is a common object in linear algebra. We say that a set S of vectors in $\mathbb{R}^n$ is a *subspace* if whenever $\vec{x}$ and $\vec{y}$ are members of S and α is a scalar, then

$$\vec{x} + \vec{y}, \qquad \text{and} \qquad \alpha\vec{x}$$

are also members of S. That is, we can add and multiply by scalars and we still land in S. So every linear combination of vectors of S is still in S. That is really what a subspace is. It is a subset where we can take linear combinations and still end up being in the subset.

Example 75. *If we let $S = \mathbb{R}^n$, then this S is a subspace of $\mathbb{R}^n$. Adding any two vectors in $\mathbb{R}^n$ gets a vector in $\mathbb{R}^n$, and so does multiplying by scalars.*

The set $S' = \{\vec{0}\}$, that is, the set of the zero vector by itself, is also a subspace of $\mathbb{R}^n$. There is only one vector in this subspace, so we only need to check for that one vector, and everything checks out: $\vec{0} + \vec{0} = \vec{0}$ and $\alpha\vec{0} = \vec{0}$.

The set S'' of all the vectors of the form (a, a) for any real number a, such as $(1, 1)$, $(3, 3)$, or $(-0.5, -0.5)$ is a subspace of $\mathbb{R}^2$. Adding two such vectors, say $(1, 1) + (3, 3) = (4, 4)$ again gets a vector of the same form, and so does multiplying by a scalar, say $8(1, 1) = (8, 8)$.

We can apply these ideas to the vectors that live inside a matrix. The span of the rows of a matrix A is called the *row space*. The row space of A and the row space of the row echelon form of A are the same, because reducing the matrix A to its row echelon form involves taking linear combinations, which will preserve the span. In the example,

$$\begin{aligned} \text{row space of } \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} &= \operatorname{span}\left\{ \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}, \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}, \begin{bmatrix} 7 & 8 & 9 \end{bmatrix} \right\} \\ &= \operatorname{span}\left\{ \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 \end{bmatrix} \right\}. \end{aligned}$$

Similarly to row space, the span of columns is called the *column space*.

$$\text{column space of } \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} \right\}.$$

So it may also be good to find the number of linearly independent columns of A. One way to do that is to find the number of linearly independent rows of A^T. It is a tremendously useful fact that the number of linearly independent columns is always the same as the number of linearly independent rows:

Theorem 14. $\operatorname{rank} A = \operatorname{rank} A^T$

In particular, to find a set of linearly independent columns we need to look at where the pivots were. If you recall above, when solving $A\vec{x} = \vec{0}$ the key was finding the pivots, any non-pivot columns corresponded to free variables. That means we can solve for the non-pivot columns in terms of the pivot columns. Let's see an example.

Example 76. *Find the linearly independent columns of the matrix*

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

Solution: We find a pivot and reduce the rows below:

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

We find the next pivot, make it one, and rinse and repeat:

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

The final matrix is the row echelon form of the matrix. Consider the pivots that we marked. The pivot columns are the first and the third column. All other columns correspond to free variables when solving $A\vec{x} = \vec{0}$, so all other columns can be solved in terms of the first and the third column. In other words

$$\text{column space of } \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix} \right\} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}.$$

We could perhaps use another pair of columns to get the same span, but the first and the third are guaranteed to work because they are pivot columns.

The discussion above could be expanded into a proof of the theorem if we wanted. As each nonzero row in the row echelon form contains a pivot, then the rank is the number of pivots, which is the same as the maximal number of linearly independent columns.

In the previous example, this means that only the first and third colums are "important" in the sense of generating the full column space as a span. We would like to have a way to talk about what these first and third columns do.

Definition 20 (Spanning set)**.** *Let S be a subspace of a vector space. The set $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_k\}$ is a* spanning set *for the subspace S if each of these vectors are in S and the span of $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_k\}$ is equal to S.*

In the context of the previous example, for the matrix

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

we know that

$$\text{column space of } \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix} \right\} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}.$$

This means that both

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix} \right\} \quad \text{and} \quad \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}$$

are spanning sets for this column space.

The idea also works in reverse. Suppose we have a bunch of column vectors and we just need to find a linearly independent set. For example, suppose we started with the vectors

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix}.$$

These vectors are not linearly independent as we saw above. In particular, the span $\vec{v}_1$ and $\vec{v}_3$ is the same as the span of all four of the vectors. So $\vec{v}_2$ and $\vec{v}_4$ can both be written as linear combinations of $\vec{v}_1$ and $\vec{v}_3$. A common thing that comes up in practice is that one gets a set of vectors whose span is the set of solutions of some problem. But perhaps we get way too many vectors, we want to simplify. For example above, all vectors in the span of $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ can be written $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \alpha_3\vec{v}_3 + \alpha_4\vec{v}_4$ for some numbers $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. But it is also true that every such vector can be written as $a\vec{v}_1 + b\vec{v}_3$ for two numbers a and b. And one has to admit, that looks much simpler. Moreover, these numbers a and b are unique. More on that later in this section.

To find this linearly independent set we simply take our vectors and form the matrix $[\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3 \ \vec{v}_4]$, that is, the matrix

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

We crank up the row-reduction machine, feed this matrix into it, and find the pivot columns and pick those. In this case, $\vec{v}_1$ and $\vec{v}_3$.

Basis and dimension

At this point, we have talked about subspaces, and two other properties of sets of vectors: linear independence and being a spanning set for a subspace. In some sense, these two properties are in opposition to each other. If I add more vectors to a set, I am more likely to become a spanning set (because I have more options for adding to get other vectors), but less likely to be independent (because there are more possibilities for a linear combination to be zero). Similarly, the reverse is true; removing vectors means the set is more likely to be linearly independent, but less likely to span a given subspace. The question then becomes if there is a sweet spot where both things are true, and that leads to the definition of a basis.

Definition 21. *If S is a subspace and we can find k linearly independent vectors in S*

$$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k,$$

such that every other vector in S is a linear combination of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$, then the set $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is called a basis *of S. In other words, S is the span of $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$. We say that S has* dimension *k, and we write*

$$\dim S = k.$$

The next theorem illustrates the main properties and classification of a basis of a vector space.

Theorem 15. *If $S \subset \mathbb{R}^n$ is a subspace and S is not the trivial subspace $\{\vec{0}\}$, then there exists a unique positive integer k (the dimension) and a (not unique) basis $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$, such that every $\vec{w}$ in S can be uniquely represented by*

$$\vec{w} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k,$$

for some scalars α_1, α_2, $\ldots$, α_k.

We should reiterate that while k is unique (a subspace cannot have two different dimensions), the set of basis vectors is not at all unique. There are lots of different bases for any given subspace. Finding just the right basis for a subspace is a large part of what one does in linear algebra. In fact, that is what we spend a lot of time on in linear differential equations, although at first glance it may not seem like that is what we are doing.

Example 77. *The standard basis*

$$\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n,$$

is a basis of $\mathbb{R}^n$ (hence the name). So as expected

$$\dim \mathbb{R}^n = n.$$

On the other hand the subspace $\{\vec{0}\}$ *is of dimension* 0.

The subspace S'' *from a previous example, that is, the set of vectors* (a,a) *is of dimension 1. One possible basis is simply* $\{(1,1)\}$, *the single vector* $(1,1)$: *every vector in* S'' *can be represented by* $a(1,1)=(a,a)$. *Similarly another possible basis would be* $\{(-1,-1)\}$. *Then the vector* (a,a) *would be represented as* $(-a)(-1,-1)$. *In this case, the subspace* S'' *has many different bases, two of which are* $\{(1,1)\}$ *and* $\{(-1,-1)\}$, *and the vector* (a,a) *has a different representation (different constant) for the different bases.*

Row and column spaces of a matrix are also examples of subspaces, as they are given as the span of vectors. We can use what we know about rank, row spaces, and column spaces from the previous section to find a basis.

Example 78. *Earlier, we considered the matrix*

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

Using row reduction to find the pivot columns, we found

$$\textit{column space of } A\left(\begin{bmatrix}1&2&3&4\\2&4&5&6\\3&6&7&8\end{bmatrix}\right)=\operatorname{span}\left\{\begin{bmatrix}1\\2\\3\end{bmatrix},\begin{bmatrix}3\\5\\7\end{bmatrix}\right\}.$$

What we did was we found the basis of the column space. The basis has two elements, and so the column space of A *is two dimensional. Notice that the rank of* A *is two.*

We would have followed the same procedure if we wanted to find the basis of the subspace X spanned by

$$\begin{bmatrix}1\\2\\3\end{bmatrix},\begin{bmatrix}2\\4\\6\end{bmatrix},\begin{bmatrix}3\\5\\7\end{bmatrix},\begin{bmatrix}4\\6\\8\end{bmatrix}.$$

We would have simply formed the matrix A with these vectors as columns and repeated the computation above. The subspace X is then the column space of A.

Example 79. *Consider the matrix*

$$L=\begin{bmatrix}1&2&0&0&3\\0&0&1&0&4\\0&0&0&1&5\end{bmatrix}$$

Conveniently, the matrix is in reduced row echelon form. The matrix is of rank 3. The column space is the span of the pivot columns, because the pivot columns always form a basis for the column space of a matrix. It is the 3-dimensional space

$$\textit{column space of } L=\operatorname{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}=\mathbb{R}^3.$$

The row space is the 3-dimensional space

$$\textit{row space of } L=\operatorname{span}\left\{\begin{bmatrix}1&2&0&0&3\end{bmatrix},\begin{bmatrix}0&0&1&0&4\end{bmatrix},\begin{bmatrix}0&0&0&1&5\end{bmatrix}\right\}.$$

As these vectors have 5 components, we think of the row space of L *as a subspace of* $\mathbb{R}^5$.

The way the dimensions worked out in the examples is not an accident. Since the number of vectors that we needed to take was always the same as the number of pivots, and the number of pivots is the rank, we get the following result.

Theorem 16. *Rank The dimension of the column space and the dimension of the row space of a matrix* A *are both equal to the rank of* A.

`linearAlgebra/practice/rankIndependence-practice1.tex`

Practice for Forced Oscillations

Why?

Exercise 438 *Compute the rank of the given matrices*

a) $\begin{bmatrix} 6 & 3 & 5 \\ 1 & 4 & 1 \\ 7 & 7 & 6 \end{bmatrix}$ b) $\begin{bmatrix} 5 & -2 & -1 \\ 3 & 0 & 6 \\ 2 & 4 & 5 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 2 & 4 & 6 \end{bmatrix}$

Exercise 439 *Compute the rank of the given matrices*

a) $\begin{bmatrix} 7 & -1 & 6 \\ 7 & 7 & 7 \\ 7 & 6 & 2 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}$ c) $\begin{bmatrix} 0 & 3 & -1 \\ 6 & 3 & 1 \\ 4 & 7 & -1 \end{bmatrix}$

Exercise 440 *For the matrices in Exercise , find a linearly independent set of row vectors that span the row space (they don't need to be rows of the matrix).*

Exercise 441 *For the matrices in Exercise , find a linearly independent set of columns that span the column space. That is, find the pivot columns of the matrices.*

Exercise 442 *For the matrices in Exercise , find a linearly independent set of row vectors that span the row space (they don't need to be rows of the matrix).*

Exercise 443 *For the matrices in Exercise , find a linearly independent set of columns that span the column space. That is, find the pivot columns of the matrices.*

Exercise 444 *Compute the rank of the matrix*

$$\begin{bmatrix} 10 & -2 & 11 & -7 \\ -5 & -2 & -5 & 5 \\ 1 & 0 & -4 & -4 \\ 1 & 2 & 2 & -1 \end{bmatrix}$$

Exercise 445 *Compute the rank of the matrix*

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

Author(s): Matthew Charnley and Jason Nowell

Exercise 446 *Find a linearly independent subset of the following vectors that has the same span.*

$$\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ -2 \\ -4 \end{bmatrix}, \quad \begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}$$

Exercise 447 *Find a linearly independent subset of the following vectors that has the same span.*

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ 1 \\ -5 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} -3 \\ 2 \\ 4 \end{bmatrix}$$

Exercise 448 *For the following sets of vectors, determine if the set is linearly independent. Then find a basis for the subspace spanned by the vectors, and find the dimension of the subspace.*

a) $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$

b) $\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$

c) $\begin{bmatrix} -4 \\ -3 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$

d) $\begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$

e) $\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

f) $\begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix}, \begin{bmatrix} -5 \\ -5 \\ -2 \end{bmatrix}$

Exercise 449 *For the following sets of vectors, determine if the set is linearly independent. Then find a basis for the subspace spanned by the vectors, and find the dimension of the subspace.*

a) $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

b) $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$

c) $\begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ -1 \\ 5 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ -4 \end{bmatrix}$

d) $\begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 4 \\ -3 \end{bmatrix}$

e) $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix}$

f) $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$

Exercise 450 *Suppose that X is the set of all the vectors of $\mathbb{R}^3$ whose third component is zero. Is X a subspace? And if so, find a basis and the dimension.*

Exercise 451 *Consider a set of 3 component vectors.*

a) *How can it be shown if these vectors are linearly independent?*

b) *Can a set of 4 of these 3 component vectors be linearly independent? Explain your answer.*

c) *Can a set of 2 of these 3 component vectors be linearly independent? Explain.*

d) *How would it be shown if these vectors make up a spanning set for all 3 component vectors?*

e) *Can 4 vectors be a spanning set? Explain.*

f) *Can 2 vectors be a spanning set? Explain.*

Exercise 452 *Consider the vectors*

$$\vec{v}_1 = \begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}.$$

Let A be the matrix with these vectors as columns and $\vec{b}$ the vector $[1\ 0\ 0]$.

a) *Compute the rank of A to determine how many of these vectors are linearly independent.*

b) *Determine if $\vec{b}$ is in the span of the given vectors by using row reduction to try to solve $A\vec{x} = \vec{b}$.*

c) *Look at the columns of the row-reduced form of A. Is $\vec{b}$ in the span of those vectors?*

d) *What do these last two parts tell you about the span of the columns of a matrix, and the span of the columns of the row-reduced matrix?*

e) *Now, build a matrix D with these vectors as rows. Row-reduce this matrix to get a matrix D_2.*

f) *Is $\vec{b}$ in the span of the rows of D_2? You can't check this in using the matrix form; instead, just brute force it based on the form of D_2. What does this potentially say about the span of the rows of D and the rows of D_2?*

Exercise 453 *Complete* Exercise *with*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} -6 \\ 2 \\ 3 \\ -1 \end{bmatrix} \quad \begin{bmatrix} -13 \\ 3 \\ 1 \\ 1 \end{bmatrix} \quad \vec{v}_4 \begin{bmatrix} 11 - 1 \\ -5 \\ -1 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Determinant

We discuss Determinant

For square matrices we define a useful quantity called the *determinant*. We define the determinant of a 1×1 matrix as the value of its only entry

$$\det\left([a]\right) \overset{\text{def}}{=} a.$$

For a 2×2 matrix we define

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \overset{\text{def}}{=} ad - bc.$$

Before defining the determinant for larger matrices, we note the meaning of the determinant. An $n \times n$ matrix gives a mapping of the n-dimensional euclidean space $\mathbb{R}^n$ to itself. In particular, a 2×2 matrix A is a mapping of the plane to itself. The determinant of A is the factor by which the area of objects changes. If we take the unit square (square of side 1) in the plane, then A takes the square to a parallelogram of area $|\det(A)|$. The sign of $\det(A)$ denotes a change of orientation (negative if the axes get flipped). For example, let

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

Then $\det(A) = 1 + 1 = 2$. Let us see where A sends the unit square with vertices $(0,0)$, $(1,0)$, $(0,1)$, and $(1,1)$. The point $(0,0)$ gets sent to $(0,0)$.

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

The image of the square is another square with vertices $(0,0)$, $(1,-1)$, $(1,1)$, and $(2,0)$. The image square has a side of length $\sqrt{2}$ and is therefore of area 2. See **Normally a reference to a previous figure goes here.**.

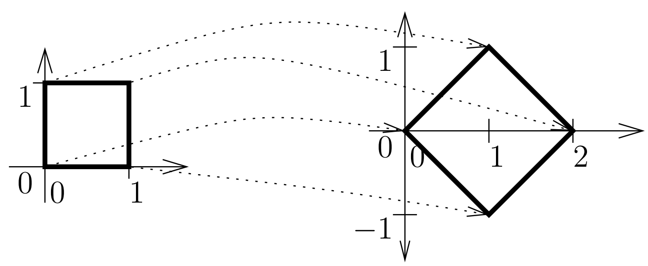


Figure 54: Image of the unit quare via the mapping A.

In general the image of a square is going to be a parallelogram. In high school geometry, you may have seen a formula for computing the area of a parallelogram with vertices $(0,0)$, (a,c), (b,d) and $(a+b, c+d)$. The area is

$$\left|\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)\right| = |ad - bc|.$$

The vertical lines above mean absolute value. The matrix $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ carries the unit square to the given parallelogram.

There are a number of ways to define the determinant for an $n \times n$ matrix. Let us use the so-called *cofactor expansion*. We define A_{ij} as the matrix A with the i^{th} row and the j^{th} column deleted. For example,

$$\text{If} \qquad A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \qquad \text{then} \qquad A_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} \qquad \text{and} \qquad A_{23} = \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix}.$$

Learning outcomes: Compute the determinant of a 2×2 matrix Use cofactor expansion to compute the determinant of larger matrices Use the determinant to make statements about invertibility or linear independence of the columns of that matrix.

Author(s): Matthew Charnley and Jason Nowell

We now define the determinant recursively

$$\det(A) \overset{\text{def}}{=} \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(A_{1j}),$$

or in other words

$$\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13}) - \cdots \begin{cases} +a_{1n}\det(A_{1n}) & \text{if } n \text{ is odd,} \\ -a_{1n}\det(A_{1n}) & \text{if } n \text{ even.} \end{cases}$$

For a 3×3 matrix, we get $\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13})$. For example,

$$\begin{aligned} \det\left(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}\right) &= 1 \cdot \det\left(\begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}\right) - 2 \cdot \det\left(\begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}\right) + 3 \cdot \det\left(\begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}\right) \\ &= 1(5 \cdot 9 - 6 \cdot 8) - 2(4 \cdot 9 - 6 \cdot 7) + 3(4 \cdot 8 - 5 \cdot 7) = 0. \end{aligned}$$

It turns out that we did not have to necessarily use the first row. That is for any i,

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

It is sometimes useful to use a row other than the first. In the following example it is more convenient to expand along the second row. Notice that for the second row we are starting with a negative sign.

$$\begin{aligned} \det\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 0 \\ 7 & 8 & 9 \end{bmatrix}\right) &= -0 \cdot \det\left(\begin{bmatrix} 2 & 3 \\ 8 & 9 \end{bmatrix}\right) + 5 \cdot \det\left(\begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix}\right) - 0 \cdot \det\left(\begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix}\right) \\ &= 0 + 5(1 \cdot 9 - 3 \cdot 7) + 0 = -60. \end{aligned}$$

Let us check if it is really the same as expanding along the first row,

$$\begin{aligned} \det\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 0 \\ 7 & 8 & 9 \end{bmatrix}\right) &= 1 \cdot \det\left(\begin{bmatrix} 5 & 0 \\ 8 & 9 \end{bmatrix}\right) - 2 \cdot \det\left(\begin{bmatrix} 0 & 0 \\ 7 & 9 \end{bmatrix}\right) + 3 \cdot \det\left(\begin{bmatrix} 0 & 5 \\ 7 & 8 \end{bmatrix}\right) \\ &= 1(5 \cdot 9 - 0 \cdot 8) - 2(0 \cdot 9 - 0 \cdot 7) + 3(0 \cdot 8 - 5 \cdot 7) = -60. \end{aligned}$$

In computing the determinant, we alternately add and subtract the determinants of the submatrices A_{ij} multiplied by a_{ij} for a fixed i and all j. The numbers $(-1)^{i+j}\det(A_{ij})$ are called *cofactors* of the matrix. And that is why this method of computing the determinant is called the *cofactor expansion*.

Similarly we do not need to expand along a row, we can expand along a column. For any j

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

A related fact is that

$$\det(A) = \det(A^T).$$

Recall that a matrix is *upper triangular* if all elements below the main diagonal are 0. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 9 \end{bmatrix}$$

is upper triangular. Similarly a *lower triangular* matrix is one where everything above the diagonal is zero. For example,

$$\begin{bmatrix} 1 & 0 & 0 \\ 4 & 5 & 0 \\ 7 & 8 & 9 \end{bmatrix}.$$

The determinant for triangular matrices is very simple to compute. Consider the lower triangular matrix. If we expand along the first row, we find that the determinant is 1 times the determinant of the lower triangular matrix $\left[\begin{smallmatrix}5&0\\8&9\end{smallmatrix}\right]$. So the deteriminant is just the product of the diagonal entries:

$$\det\left(\begin{bmatrix}1&0&0\\4&5&0\\7&8&9\end{bmatrix}\right) = 1\cdot 5\cdot 9 = 45.$$

Similarly for upper triangular matrices

$$\det\left(\begin{bmatrix}1&2&3\\0&5&6\\0&0&9\end{bmatrix}\right) = 1\cdot 5\cdot 9 = 45.$$

In general, if A is triangular, then

$$\det(A) = a_{11}a_{22}\cdots a_{nn}.$$

If A is diagonal, then it is also triangular (upper and lower), so same formula applies. For example,

$$\det\left(\begin{bmatrix}2&0&0\\0&3&0\\0&0&5\end{bmatrix}\right) = 2\cdot 3\cdot 5 = 30.$$

In particular, the identity matrix I is diagonal, and the diagonal entries are all 1. Thus,

$$\det(I) = 1.$$

Another way that we can compute determinants is by using row reduction. Since the row echelon form is a diagonal matrix, this will make it easy to compute the determinant using the product of the diagonal entries. However, we need to know how the determinant is affected by elementary row operations.

Theorem 17 (Properties of the Determinant)**.** *Let A be a square $n \times n$ matrix.*

(a) *If B obtained from A by interchanging two rows (or two columns) of A, then* $\det(B) = -\det(A)$.

(b) *If B is obtained from A by multiplying a row of column by the number r, then* $\det(B) = r\det(A)$.

(c) *If B is obtained from A by multiplying a row (or column) by a non-zero number r and adding the result to another row, then* $\det(B) = \det(A)$.

Proof The proof of each of these facts comes from the cofactor expansion of the determinant.

(a) Assume that B is obtained by interchanging the first and second row of A. We will use cofactor expansion along the first row to find the determinant of A, and the second row for the determinant of B. We get that

$$\det(A) = \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det(A_{1j})$$

and

$$\det(B) = \sum_{j=1}^{n}(-1)^{2+j}b_{2j}\det(B_{2j}).$$

However, since the second row of B is the first row of A, we know that $b_{2j} = a_{1j}$ for all j. In addition, this swap means that we also have that $B_{2j} = A_{1j}$ for each of the cofactors in this expansion. All of these cofactor matrices are made up of the second through last rows of A, with the appropriate columns removed at each step.

Therefore, the only difference between these two formulas is that the A formula starts with $(-1)^{1+j}$ and the B formula starts with $(-1)^{2+j}$. Thus, $\det(B)$ will have an additional factor of -1 in it, giving the desired result.

The exact same process works for swapping any two adjacent rows of the matrix, giving that this also provides a -1 in the computation of the determinant. For non-adjacent rows, we use the fact that to any swap of non-adjacent rows of a matrix requires an *odd* number of adjacent row swaps. For example, if we want to swap rows 1 and 3, we can swap row 1 with row 2, then row 2 with row 3, and finally swap row 1 with row 2 again. This will put the first row in the third spot and the third row up in the first slot. Since each of these adjacent switches adds a minus sign, doing an odd number of switches still results in adding a single minus sign to the computation of the determinant.

(b) Assume that we want to multiply the kth row of A by the number r to get B. We use cofactor expansion along this same kth row to find the determinant of each matrix. We get that

$$\det(A) = \sum_{j=1}^{n}(-1)^{k+j}a_{kj}\det(A_{kj})$$

and

$$\det(B) = \sum_{j=1}^{n}(-1)^{k+j}b_{kj}\det(B_{kj}) = \sum_{j=1}^{n}(-1)^{k+j}ra_{kj}\det(B_{kj}).$$

However, the minor B_{kj} ignores the kth row of the matrix B, so the minors are identical to those of A. Thus, we have that

$$\det(B) = \sum_{j=1}^{n}(-1)^{k+j}ra_{kj}\det(B_{kj}) = r\sum_{j=1}^{n}(-1)^{k+j}a_{kj}\det(A_{kj}) = r\det(A).$$

(c) Assume that B is formed by adding r copies of the kth row of A to the ith row. Since the ith row is the one being changed, we will use cofactor expansion there to compute each determinant. We get that

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

and

$$\det(B) = \sum_{j=1}^{n}(-1)^{i+j}b_{ij}\det(B_{ij}) = \sum(-1)^{i+j}(a_{ij}+ra_{kj})\det(A_{ij})$$

where we have replaced the minors of B by the minors of A because they ignore the ith row, which is the only thing that has changed. We can now split the determinant of B into two parts

$$\sum(-1)^{i+j}(a_{ij}+ra_{kj})\det(A_{ij}) = \sum(-1)^{i+j}a_{ij}\det(A_{ij}) + \sum(-1)^{i+j}ra_{kj}\det(A_{ij}).$$

The first of these is the determinant of the matrix A. The second is the determinant of a new matrix that we will call C. C is the same as the matrix A, except that we have replaced the ith row of A by r times the kth row of A. Thus, the ith row of this matrix C is a multiple of the kth row. This means that the rows of C are not linearly independent. By Theorem 20 coming up later (don't worry, it does not depend on this result), this tells us that the determinant of C is zero. Therefore

$$\det(B) = \det(A) + \det(C) = \det(A)$$

so this operation does not change the determinant of the matrix.

■

These correspond to the three elementary row operations that we use to row reduce matrices. In order to use this to compute determinants, we need to keep track of each of these operations and how the determinant changes at each step.

Example 80. *Compute the determinant of the matrix*

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

using row reduction.

Solution: We will go through the process of row reduction to find the determinant. We need to keep track of each time that we swap rows (to add a minus sign) and that we multiply a row by a constant (to factor in that constant). Throughout this process, we will use A to refer to the initial matrix

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

and M will refer to wherever we are in the process. So we will start by dividing the first row of the matrix by -4

$$\begin{bmatrix} -4 & -2 & 4 \\ -3 & -3 & 2 \\ -2 & -3 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1/2 & -1 \\ -3 & -3 & 2 \\ -2 & -3 & 1 \end{bmatrix}.$$

Since we divided by -4, Theorem 17 tells us that

$$\det(M) = -\frac{1}{4}\det(A).$$

The next step of row reduction will be to use the 1 in the top left to cancel out the -3 and -2 below it. Part (c) in Theorem 17 says that this doesn't change the determinant. Therefore, the row reduction gives

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

and we still have that

$$\det(M) = -\frac{1}{4}\det(A).$$

Next, we will multiply row 2 by $-\frac{2}{3}$, which gives

$$\begin{bmatrix} 1 & 1/2 & -1 \\ 0 & -3/2 & -1 \\ 0 & -2 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1/2 & -1 \\ 0 & 1 & 2/3 \\ 0 & -2 & -1 \end{bmatrix}.$$

Adding this in to our previous steps using Theorem 17, we get that

$$\det(M) = \left(-\frac{2}{3}\right)\left(-\frac{1}{4}\right)\det(A).$$

Finally, we add two copies of row 2 to row 3, which does not change the determinant and gives the matrix

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

with

$$\det(M) = \left(-\frac{2}{3}\right)\left(-\frac{1}{4}\right)\det(A).$$

We can rearrange this expression to say that

$$\det(A) = 6\det(M)$$

and we can easily compute that $\det(M) = \frac{1}{3}$ by multiplying the diagonal entries. Thus, we have that $\det(A) = 2$.

Exercise 454 *Compute* $\det(A)$ *using cofactor expansion and show that you get the same answer.*

The determinant is telling you how geometric objects scale. If B doubles the sizes of geometric objects and A triples them, then AB (which applies B to an object and then it applies A) should make size go up by a factor of 6. This is true in general:

Theorem 18.

$$\det(AB) = \det(A)\det(B).$$

This property is one of the most useful, and it is employed often to actually compute determinants. A particularly interesting consequence is to note what it means for existence of inverses. Take A and B to be inverses, that is $AB = I$. Then

$$\det(A)\det(B) = \det(AB) = \det(I) = 1.$$

Neither $\det(A)$ nor $\det(B)$ can be zero. This fact is an extremely useful property of the determinant, and one which is used often in this book:

Theorem 19. *An $n \times n$ matrix A is invertible if and only if* $\det(A) \neq 0$.

In fact, $\det(A^{-1})\det(A) = 1$ says that

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

So we know what the determinant of A^{-1} is without computing A^{-1}.

Let us return to the formula for the inverse of a 2×2 matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Notice the determinant of the matrix $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ in the denominator of the fraction. The formula only works if the determinant is nonzero, otherwise we are dividing by zero.

A common notation for the determinant is a pair of vertical lines:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right).$$

Personally, I find this notation confusing as vertical lines usually mean a positive quantity, while determinants can be negative. Also think about how to write the absolute value of a determinant. This notation is not used in this book.

With this discussion of determinants complete, we can now state a major theorem from linear algebra that will help us here and when we get back to solving differential equations using this linear algebra. In a full course on linear algebra, this theorem would be covered in full detail, including all of the proofs. For this introduction, we give some idea as to why everything is true here, but not all of the details.

Note: This is an example of an *equivalence* theorem, which is fairly common in mathematics. It means that if any one of the statements are true, then we know that all of the others are true as well. It means it's harder to prove, but once we have such a theorem, it is very powerful in how we can use it going forward.

Theorem 20. *Let A be an $n \times n$ matrix. The following are equivalent:*

(a) *A is invertible.*

(b) $\det(A) \neq 0$.

(c) *There is a unique solution to $A\vec{x} = \vec{b}$ for every vector $\vec{b}$.*

(d) *The only solution to $A\vec{x} = \vec{0}$ is $\vec{x} = \vec{0}$.*

(e) *The reduced row echelon form of A is I_n, the identity matrix.*

(f) *The columns of A are linearly independent.*

Proof Why is all of this true? For (a) and (b), we have Theorem 19 to say that they are equivalent. For (c), if A is invertible, then the unique solution to $A\vec{x} = \vec{b}$ is $\vec{x} = A^{-1}\vec{b}$. If we take $\vec{b} = \vec{0}$ here, we get (d), that the solution is $\vec{x} = A^{-1}\vec{0} = \vec{0}$. This means that reducing the system of equations $A\vec{x} = 0$ gives $x_1 = 0$, $x_2 = 0$, ..., $x_n = 0$, which means the reduced row echelon form of A is just the identity matrix, which is (e). Finally, this means that every column is a pivot column, so that all of the columns are linearly independent, giving (f). ■

This is a massive theorem that forms most of the backbone of linear algebra. We will only be using a few parts of it later and we can also add some other parts to it with different definitions from linear algebra, but since we have seen all of these components, it is nice to see them all put together into one complete statement.

`linearAlgebra/practice/determinant-practice1.tex`

Practice for Determinants

Why?

Exercise 455 *Compute the determinant of the following matrices:*

a) $\begin{bmatrix} 3 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$ c) $\begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}$ d) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$

e) $\begin{bmatrix} 2 & 1 & 0 \\ -2 & 7 & -3 \\ 0 & 2 & 0 \end{bmatrix}$ f) $\begin{bmatrix} 2 & 1 & 3 \\ 8 & 6 & 3 \\ 7 & 9 & 7 \end{bmatrix}$ g) $\begin{bmatrix} 0 & 2 & 5 & 7 \\ 0 & 0 & 2 & -3 \\ 3 & 4 & 5 & 7 \\ 0 & 0 & 2 & 4 \end{bmatrix}$ h) $\begin{bmatrix} 0 & 1 & 2 & 0 \\ 1 & 1 & -1 & 2 \\ 1 & 1 & 2 & 1 \\ 2 & -1 & -2 & 3 \end{bmatrix}$

Exercise 456 *% Compute the determinant of the following matrices:*

a) $\begin{bmatrix} -2 \end{bmatrix}$ b) $\begin{bmatrix} 2 & -2 \\ 1 & 3 \end{bmatrix}$ c) $\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ d) $\begin{bmatrix} 2 & 9 & -11 \\ 0 & -1 & 5 \\ 0 & 0 & 3 \end{bmatrix}$

e) $\begin{bmatrix} 2 & 1 & 0 \\ -2 & 7 & 3 \\ 1 & 1 & 0 \end{bmatrix}$ f) $\begin{bmatrix} 5 & 1 & 3 \\ 4 & 1 & 1 \\ 4 & 5 & 1 \end{bmatrix}$ g) $\begin{bmatrix} 3 & 2 & 5 & 7 \\ 0 & 0 & 2 & 0 \\ 0 & 4 & 5 & 0 \\ 2 & 1 & 2 & 4 \end{bmatrix}$ h) $\begin{bmatrix} 0 & 2 & 1 & 0 \\ 1 & 2 & -3 & 4 \\ 5 & 6 & -7 & 8 \\ 1 & 2 & 3 & -2 \end{bmatrix}$

Exercise 457 *For which x are the following matrices singular (not invertible).*

a) $\begin{bmatrix} 2 & 3 \\ 2 & x \end{bmatrix}$ b) $\begin{bmatrix} 2 & x \\ 1 & 2 \end{bmatrix}$ c) $\begin{bmatrix} x & 1 \\ 4 & x \end{bmatrix}$ d) $\begin{bmatrix} x & 0 & 1 \\ 1 & 4 & 2 \\ 1 & 6 & 2 \end{bmatrix}$

Exercise 458 *% For which x are the following matrices singular (not invertible).*

a) $\begin{bmatrix} 1 & 3 \\ 1 & x \end{bmatrix}$ b) $\begin{bmatrix} 3 & x \\ 1 & 3 \end{bmatrix}$ c) $\begin{bmatrix} x & 3 \\ 3 & x \end{bmatrix}$ d) $\begin{bmatrix} x & 1 & 0 \\ 1 & 4 & 0 \\ 1 & 6 & 2 \end{bmatrix}$

Exercise 459 *% Consider the matrix*

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

a) *Compute the determinant of A using cofactor expansion along row 1.*

b) *Compute the determinant of A using cofactor expansion along column 2.*

c) *Compute the determinant using row reduction.*

Author(s): Matthew Charnley and Jason Nowell

Exercise 460 *% Consider the matrix*

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

a) *Compute the determinant of A using cofactor expansion along row 1.*

b) *Compute the determinant of A using cofactor expansion along column 3.*

c) *Compute the determinant using row reduction.*

Exercise 461 *% Consider the matrix*

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

a) *Compute the determinant of A using cofactor expansion along row 1.*

b) *Compute the determinant of A using cofactor expansion along column 4.*

c) *Compute the determinant using row reduction.*

Exercise 462 *Is the matrix A below invertible? How do you know?*

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

Exercise 463 *% Compute the determinant of the matrix*

$$A = \begin{bmatrix} 5 & 4 & 3 \\ -4 & -3 & -4 \\ -5 & -5 & 4 \end{bmatrix}$$

using row reduction. What does this say about the solutions to $A\vec{x} = 0$?

Exercise 464 *% Compute the determinant of the matrix*

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

using row reduction. What does this say about the columns of A?

Exercise 465 *% Compute the determinant of the matrix*

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

using row reduction. What does this say about the solutions to $A\vec{x} = \left[\begin{smallmatrix} 1 \\ 0 \\ -1 \\ 1 \end{smallmatrix}\right]$.

Exercise 466 *Compute*

$$\det\left(\begin{bmatrix} 2 & 1 & 2 & 3 \\ 0 & 8 & 6 & 5 \\ 0 & 0 & 3 & 9 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1}\right)$$

without computing the inverse.

Exercise 467 *% Compute*

$$\det\left(\begin{bmatrix} 3 & 4 & 7 & 12 \\ 0 & -1 & 9 & -8 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & 0 & 2 \end{bmatrix}^{-1}\right)$$

without computing the inverse.

Exercise 468 *Suppose*

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 7 & \pi & 1 & 0 \\ 2^8 & 5 & -99 & 1 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 5 & 9 & 1 & -\sin(1) \\ 0 & 1 & 88 & -1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Let $A = LU$*. Compute* $\det(A)$ *in a simple way, without computing what is* A*. Hint: First read off* $\det(L)$ *and* $\det(U)$*.*

Exercise 469 *Consider the linear mapping from* $\mathbb{R}^2$ *to* $\mathbb{R}^2$ *given by the matrix* $A = \left[\begin{smallmatrix} 1 & x \\ 2 & 1 \end{smallmatrix}\right]$ *for some number* x*. You wish to make* A *such that it doubles the area of every geometric figure. What are the possibilities for* x *(there are two answers).*

Exercise 470 *% Consider the matrix*

$$A(x) = \begin{bmatrix} 1 & 2 \\ 1 & x \end{bmatrix}^{-1}$$

as a function of the variable x*.*

a) *Find all the* x *so that* $A(x)$ *and the matrix inverse* $A(x)^{-1}$ *have only integer entries (no fractions). Note that there are two answers.*

b) *Find all the* x *so that the matrix inverse* $A(x)^{-1}$ *has only integer entries (no fractions). (You should get more answers here than the previous part.)*

Exercise 471 *Suppose A and S are $n \times n$ matrices, and S is invertible. Suppose that $\det(A) = 3$. Compute $\det(S^{-1}AS)$ and $\det(SAS^{-1})$. Justify your answer using the theorems in this section.*

Exercise 472 *Let A be an $n \times n$ matrix such that $\det(A) = 1$. Compute $\det(xA)$ given a number x. Hint: First try computing $\det(xI)$, then note that $xA = (xI)A$.*

Eigenvalues and Eigenvectors

We discuss Eigenvalues and Eigenvectors

Consider the matrix

$$A = \begin{bmatrix} 7 & -8 \\ 3 & -3 \end{bmatrix}.$$

We can compute a few operations with this matrix. For instance

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 & -8 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

and

$$A \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 & -8 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}.$$

This last computation is fairly interesting, because the result we get is the same as 3 times the original vector. However, the matrix A does not multiply every vector by 3, as seen in the first example and the fact that

$$A \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

so A actually preserves this vector, multiplying it by 1. So, these vectors, $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$, and numbers, 3 and 1, are somehow special for this matrix A. With this information, we want to define these vectors as *eigenvectors* and numbers as *eigenvalues* of the matrix A.

Definition 22. *For a square matrix A, we say that non-zero vector $\vec{v}$ is an* eigenvector *of the matrix A if there exists a number λ so that*

$$A\vec{v} = \lambda\vec{v}.$$

In this case, we say that λ is an eigenvalue *of A and it is the* corresponding eigenvalue *for the eigenvector $\vec{v}$.*

Thus, we can say that, for the matrix

$$A = \begin{bmatrix} 7 & -8 \\ 3 & -3 \end{bmatrix},$$

we see that $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector with corresponding eigenvalue 3, and that $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ is an eigenvector with corresponding eigenvalue 1.

Why are these important? It turns out that these eigenvalues and eigenvectors characterize the behavior of the matrix A. For example, if we wanted to figure out what happens when A is applied to the vector $\begin{bmatrix} 6 \\ 4 \end{bmatrix}$, we can figure this out as

$$\begin{aligned} A \begin{bmatrix} 6 \\ 4 \end{bmatrix} &= A \left(\begin{bmatrix} 4 \\ 3 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right) \\ &= A \begin{bmatrix} 4 \\ 3 \end{bmatrix} + A \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ 3 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 6 \end{bmatrix} \end{aligned}$$

In addition, eigenvectors determine directions in which multiplying by the matrix A behaves just like scalar multiplication. This idea will be very important for our understanding of systems of differential equations, because we have already seen how to solve a scalar first order equation way back in § and § .

Learning outcomes: Find the eigenvalues and eigenvectors of a matrix Use complex numbers to find eigenvalues and eigenvectors if necessary Identify the algebraic and geometric multiplicity of an eigenvalue to determine if it is defective.

Author(s): Matthew Charnley and Jason Nowell

Finding Eigenvalues and Eigenvectors

Since eigenvalues and eigenvectors are so important, we want to know how to find them. To do this, we are looking for a number λ and a non-zero vector $\vec{v}$ so that

$$A\vec{v} = \lambda\vec{v}.$$

We can rewrite this as

$$A\vec{v} - \lambda\vec{v} = 0$$

or, using the identity matrix,

$$(A - \lambda I)\vec{v} = 0.$$

This means that we are looking for a non-zero solution to a homogeneous vector equation of the form $B\vec{v} = 0$. This is where all of our linear algebra theory comes into play.

Theorem 20 tells us that, combining parts (b) and (d), that there is a non-zero solution to $(A - \lambda I)\vec{v} = 0$ if and only if the determinant of the matrix $A - \lambda I$ is zero. Therefore, we can compute this determinant, find the values of λ so that $\det(A - \lambda I) = 0$, and these will give us our eigenvalues. Let's see an example of what this looks like.

Example 81. *Compute* $\det(A - \lambda I)$ *for the matrix*

$$A = \begin{bmatrix} 7 & -8 \\ 3 & -3 \end{bmatrix}.$$

Solution: For this matrix, we have that

$$A - \lambda I = \begin{bmatrix} 7 & -8 \\ 3 & -3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7-\lambda & -8 \\ 3 & -3-\lambda \end{bmatrix}.$$

Thus

$$\begin{aligned} \det(A - \lambda I) &= \det\left(\begin{bmatrix} 7-\lambda & -8 \\ 3 & -3-\lambda \end{bmatrix}\right) \\ &= (7-\lambda)(-3-\lambda) - (-8)(3) = \lambda^2 + 3\lambda - 7\lambda - 21 + 24 \\ &= \lambda^2 - 4\lambda + 3 \end{aligned}$$

If we were looking for eigenvalues here, we could then set this equal to zero, getting that

$$0 = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3)$$

so that the eigenvalues are 1 and 3.

In this case, we saw that computing $\det(A - \lambda I)$ for this case, we ended up with a quadratic polynomial, so it was easy to find the eigenvalues. Thankfully, no matter the size of the matrix, we will always get a polynomial here.

Definition 23. *For a matrix A, the expression* $\det(A - \lambda I)$ *is called the* characteristic polynomial *of the matrix. It will always be a polynomial, and for A an $n \times n$ matrix, it will be a degree n polynomial.*

This explains why we got a quadratic polynomial for the 2×2 matrix A. Therefore, for a matrix A, the roots of the characteristic polynomial are the eigenvalues of A.

Once we have the eigenvalues, we can use them to find the eigenvectors. As with how we started this discussion, we are looking for a non-zero vector $\vec{v}$ so that

$$(A - \lambda I)\vec{v} = 0,$$

and we know the value of λ. Therefore, we can set up a system of equations that correponds to

$$(A - \lambda I)\vec{v} = 0$$

and solve it for the components of the eigenvector.

Example 82. *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 7 & -8 \\ 3 & -3 \end{bmatrix}.$$

Solution: The previous example shows that the eigenvalues for this matrix are 1 and 3. For the eigenvalue 1, we want to find a non-zero solution to $(A-I)\vec{v}=0$, which means we want to solve for

$$(A-I)\vec{v}=\begin{bmatrix}7-1 & -8\\ 3 & -3-1\end{bmatrix}\vec{v}=\begin{bmatrix}6 & -8\\ 3 & -4\end{bmatrix}\vec{v}=0.$$

Writing the vector $\vec{v}$ as $\begin{bmatrix}v_1\\ v_2\end{bmatrix}$, this system of equations becomes

$$\begin{aligned}6v_1-8v_2&=0\\ 3v_1-4v_2&=0\end{aligned}.$$

Since the second equation is two times the first one, these equations are redundant, so we only need to satisfy $3v_1-4v_2=0$. We can do this by choosing $v_1=4$ and $v_2=3$, which gives that for $\lambda=1$, a corresponding eigenvector is $\begin{bmatrix}4\\ 3\end{bmatrix}$.

We can follow the same process for the eigenvalue 3. For this, we want to find a non-zero solution to $(A-3I)\vec{v}=0$, which means that we want to solve

$$(A-3I)\vec{v}=\begin{bmatrix}7-3 & -8\\ 3 & -3-3\end{bmatrix}\vec{v}=\begin{bmatrix}4 & -8\\ 3 & -6\end{bmatrix}\vec{v}=0.$$

Writing the vector $\vec{v}$ as $\begin{bmatrix}v_1\\ v_2\end{bmatrix}$, we get the two equations

$$\begin{aligned}4v_1-8v_2&=0\\ 3v_1-6v_2&=0\end{aligned}.$$

As before, these two equations are the same, since they are both a multiple of $v_1-2v_2=0$. Therefore, we just need to find a solution to that previous equation, which can be done with $v_1=2$ and $v_2=1$. Therefore, an eigenvector for eigenvalue 3 is $\begin{bmatrix}2\\ 1\end{bmatrix}$.

This example illustrates the standard process that is always used to find eigenvalues and eigenvectors of matrices: find the characteristic polynomial, get the roots of this polynomial, and use each of these eigenvalues to set up a system of equations for the components of each eigenvector. In addition, the equations that we get from this system will always be redundant if we have found the eigenvalue correctly. Since $\det(A-\lambda I)=0$, we know that the rows of the matrix $A-\lambda I$ are not linearly independent, and so the row-echelon form of $A-\lambda I$ must have a zero row in it. This process works for any size matrix, but it becomes harder to find the roots of this polynomial when it is higher degree.

Example 83. *Find the eigenvalues and eigenvectors of the matrix*

$$A=\begin{bmatrix}1 & 6 & 0\\ 9 & -4 & 10\\ 2 & -6 & 3\end{bmatrix}.$$

Solution: We start by hunting for eigenvalues by taking the determinant of $A-\lambda I$, which will require the cofactor expansion in order to solve.

$$\begin{aligned}\det(A-\lambda I)&=\det\left(\begin{bmatrix}1-\lambda & 6 & 0\\ 9 & -4-\lambda & 10\\ 2 & -6 & 3-\lambda\end{bmatrix}\right)\\ &=(1-\lambda)\det\left(\begin{bmatrix}-4-\lambda & 10\\ -6 & 3-\lambda\end{bmatrix}\right)-6\det\left(\begin{bmatrix}9 & 10\\ 2 & 3-\lambda\end{bmatrix}\right)\\ &=(1-\lambda)((-4-\lambda)(3-\lambda)+60)-6(9(3-\lambda)-20)\\ &=(1-\lambda)(\lambda^2+4\lambda-3\lambda-12+60)-6(27-9\lambda-20)\\ &=(1-\lambda)(\lambda^2+\lambda+48)-42+54\lambda\\ &=\lambda^2+\lambda+48-\lambda^3-\lambda^2-48\lambda-42+54\lambda\\ &=-\lambda^3+7\lambda+6\end{aligned}$$

We need to look for the roots of this polynomial. There's no nice way to factor this right away, so we need to start guessing roots. We know that the root must be a factor of 6. If we try $\lambda = 1$, we get

$$-1 + 7 + 6 = 12 \neq 0$$

so that one doesn't work. Plugging in $\lambda = -1$, we get

$$-(-1)^3 - 7 + 6 = 1 - 7 + 6 = 0$$

so this is a root, meaing that $\lambda + 1$ is a factor of the characteristic polynomial. We can then use polynomial long division to get that

$$-\lambda^3 + 7\lambda + 6 = (\lambda+1)(-\lambda^2 + \lambda + 6) = -(\lambda+1)(\lambda^2 - \lambda - 6)$$

and the quadratic term here factors as $(\lambda - 3)(\lambda + 2)$. Thus, the characteristic polynomial of this matrix is

$$(\lambda+1)(\lambda-3)(\lambda+2)$$

so the eigenvalues are -1, 3, and -2.

For the eigenvalue -1, the eigenvector must satisfy

$$(A + I)\vec{v} = \vec{0}$$

which we can write as

$$\begin{bmatrix} 2 & 6 & 0 \\ 9 & -3 & 10 \\ 2 & -6 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \vec{0}.$$

To solve this, we row-reduce the coefficient matrix.

$$\begin{aligned} \begin{bmatrix} 2 & 6 & 0 \\ 9 & -3 & 10 \\ 2 & -6 & 4 \end{bmatrix} &\to \begin{bmatrix} 1 & 3 & 0 \\ 9 & -3 & 10 \\ 2 & -6 & 4 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & 3 & 0 \\ 0 & -30 & 10 \\ 0 & -12 & 4 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & 3 & 0 \\ 0 & -3 & 1 \\ 0 & -12 & 4 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & 3 & 0 \\ 0 & -3 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Therefore, the eigenvector must satsify $v_1 + 3v_2 = 0$ and $-3v_2 + v_3 = 0$. We need to pick any non-zero set of numbers that solves these equations. For example, we could pick $v_2 = 1$ to get that we need $v_1 = -3$ and $v_3 = 3$. This gives an eigenvector of

$$\begin{bmatrix} -3 \\ 1 \\ 3 \end{bmatrix}.$$

For the eigenvalue 3, the eigenvector must satisfy

$$\begin{bmatrix} -2 & 6 & 0 \\ 9 & -7 & 10 \\ 2 & -6 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \vec{0}.$$

Row reduction gives

$$\begin{aligned} \begin{bmatrix} -2 & 6 & 0 \\ 9 & -7 & 10 \\ 2 & -6 & 0 \end{bmatrix} &\to \begin{bmatrix} 1 & -3 & 0 \\ 9 & -7 & 10 \\ 2 & -6 & 0 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & -3 & 0 \\ 0 & 20 & 10 \\ 0 & 0 & 0 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & -3 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

which means that the eigenvector must satisfy $v_1 - 3v_2 = 0$ and $2v_2 + v_3 = 0$. Again, choosing $v_2 = 1$ gives that we want $v_1 = 3$ and $v_3 = -2$. Therefore, a corresponding eigenvector here is

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

For the eigenvalue -2, the eigenvector must satisfy

$$\begin{bmatrix} 3 & 6 & 0 \\ 9 & -2 & 10 \\ 2 & -6 & 5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \vec{0}$$

where we can row reduce the coefficient matrix.

$$\begin{aligned} \begin{bmatrix} 3 & 6 & 0 \\ 9 & -2 & 10 \\ 2 & -6 & 5 \end{bmatrix} &\to \begin{bmatrix} 1 & 2 & 0 \\ 9 & -2 & 10 \\ 2 & -6 & 5 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & 2 & 0 \\ 0 & -20 & 10 \\ 0 & -10 & 5 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \\ 0 & -10 & 5 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}.$$

Therefore, the eigenvector must satisfy $v_1 + 2v_2 = 0$ and $-2v_2 + v_3 = 0$. Picking $v_2 = 1$ again gives that we want $v_1 = -2$ and $v_3 = 2$. Therefore, an eigenvector with eigenvalue -2 is

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

Real Eigenvalues

Since eigenvalues come from finding the roots of a polynomial, there are a few different situations that can arise in terms of these eigenvalues. If we take a quadratic polynomial, there are three options for the two roots.

- Two real and different roots,
- Two complex roots in a conjugate pair, or
- One double (repeated) root.

The same is true for eigenvalues, they are either all real and distinct, there are some that appear in complex conjugate pairs, or there are some repeated eigenvalues. The easiest of these cases is when the characteristic polynomial has all real and distinct eigenvalues.

In this case, we get a very nice result. We know that for each eigenvalue, there will always be at least one eigenvector, otherwise it wouldn't be an eigenvalue. If the matrix A is an $n \times n$ matrix, then the characteristic polynomial is a degree n polynomial, which will have n distinct roots by our assumption. Each of these will have a corresponding eigenvector, giving us n eigenvectors as well. A more involved result tells us that eigenvectors for different eigenvalues are always linearly independent. Therefore, we get n vectors in $\mathbb{R}^n$, that are linearly independent, and so they are a basis. This gives the following result.

Theorem 21. *Let A be an $n \times n$ matrix. Assume that the characteristic polynomial of A has all real and distinct roots, namely that*

$$\det(A - \lambda I) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

for $\lambda_1, ..., \lambda_n$ the distinct real eigenvalues. Then there exist vectors $\vec{v}_1$, ..., $\vec{v}_n$ such that $\vec{v}_i$ is an eigenvector for eigenvalue λ_i and $\{\vec{v}_1, ..., \vec{v}_n\}$ form a basis of $\mathbb{R}^n$.

To reference, look at the previous example. We found three distinct real eigenvalues of -1, 3, and -2. For these eigenvalues, we had eigenvectors

$$-1 \rightarrow \begin{bmatrix} -3 \\ 1 \\ 3 \end{bmatrix} \quad 3 \rightarrow \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix} \quad -2 \rightarrow \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}.$$

These three vectors are linearly independent (check this!) and since they are three component vectors, the space has dimension 3, and so 3 linearly independent vectors must make up a basis. This is useful to know for now, but will be critical when we want to use this information to solve systems of differential equations later.

Complex Eigenvalues

When the matrix has complex eigenvalues, the process is very similar to before. However, the eigenvector will necessarily also be complex, that is, some of the components of this vector will be complex numbers. Let's illustrate this with an example.

Example 84. *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 3 & -8 \\ 5 & -9 \end{bmatrix}.$$

Solution: We first look for the eigenvalues using the characteristic polynomial of A.

$$\begin{aligned} \det(A - \lambda I) &= \det\left(\begin{bmatrix} 3 - \lambda & -8 \\ 5 & -9 - \lambda \end{bmatrix}\right) \\ &= (3 - \lambda)(-9 - \lambda) + 40 \\ &= \lambda^2 + 9\lambda - 3\lambda - 27 + 40 \\ &= \lambda^2 + 6\lambda + 13 \end{aligned}.$$

This quadratic does not factor, so we use the quadratic formula to find that

$$\lambda = \frac{-6 \pm \sqrt{6^2 - 4 \cdot 13}}{2} = \frac{-6 \pm \sqrt{-16}}{2} = -3 \pm 2i$$

so that we have complex eigenvalues.

We now look for the eigenvectors in the same way as in the real case. If we take the eigenvalue $-3 + 2i$, then such an eigenvector must satisfy

$$(A - (-3 + 2i)I)\vec{v} = \vec{0}.$$

This means that

$$\begin{bmatrix} 3 - (-3 + 2i) & -8 \\ 5 & -9 - (-3 + 2i) \end{bmatrix} \vec{v} = \begin{bmatrix} 6 - 2i & -8 \\ 5 & -6 - 2i \end{bmatrix} \vec{v} = \vec{0}.$$

These two equations should be redundant, and to verify that, we will multiply the top row by $6 + 2i$ in row reduction to get

$$\begin{aligned} \begin{bmatrix} 6 - 2i & -8 \\ 5 & -6 - 2i \end{bmatrix} &\rightarrow \begin{bmatrix} (6 - 2i)(6 + 2i) & -8(6 + 2i) \\ 5 & -6 - 2i \end{bmatrix} \\ &\rightarrow \begin{bmatrix} 40 & -48 - 16i \\ 5 & -6 - 2i \end{bmatrix} \end{aligned}$$

and from this, we can see that the top row is 8 times the bottom one, so they are redundant. Thus, an eigenvector must satisfy

$$5v_1 - (6 + 2i)v_2 = 0$$

and we can pick any non-zero numbers that satisfy this. One simple way to do this is by switching the coefficients, so that $v_1 = 6 + 2i$ and $v_2 = 5$. Therefore, an eigenvector that we get is

$$\begin{bmatrix} 6+2i \\ 5 \end{bmatrix}.$$

Now, we can take the other eigenvalue, $-3 - 2i$. The process is the same, so that the vector must satisfy

$$\begin{bmatrix} 3-(-3-2i) & -8 \\ 5 & -9-(-3-2i) \end{bmatrix} \vec{v} = \begin{bmatrix} 6+2i & -8 \\ 5 & -6+2i \end{bmatrix} \vec{v} = \vec{0}.$$

To check redundancy again, we multiply the top row by $6 - 2i$ to get

$$\begin{aligned} \begin{bmatrix} 6+2i & -8 \\ 5 & -6+2i \end{bmatrix} &\to \begin{bmatrix} (6+2i)(6-2i) & -8(6-2i) \\ 5 & -6+2i \end{bmatrix} \\ &\to \begin{bmatrix} 40 & -48+16i \\ 5 & -6+2i \end{bmatrix} \end{aligned}$$

and again, the first equation is 8 times the second one. Thus, the eigenvector will need to satisfy

$$5v_1 - (6-2i)v_2 = 0$$

which can be done by picking $v_1 = 6 - 2i$ and $v_2 = 5$, giving an eigenvector of

$$\begin{bmatrix} 6-2i \\ 5 \end{bmatrix}.$$

The process here is the same as it was in the real case, except that now all of the equations are complex equations. In particular, the "redundancy" that we expect to see between the equations will likely be via a complex multiple. The easiest way to verify that these equations are redundant is by multiplying the first entry in each row by its complex conjugate. This is because, if we have the complex number $a + bi$, multiplying this by $a - bi$ gives

$$(a+bi)(a-bi) = a^2 + abi - abi - b^2i^2 = a^2 + b^2$$

which is now a real number. This will make it easier to compare the two equations to make sure that they are redundant, and that the eigenvalue was found correctly.

Example 85. *Find the eigenvalues and eigenvectors of the matrix*

$$A = \begin{bmatrix} 1 & 9 & 6 \\ 0 & 1 & 6 \\ 0 & -3 & -5 \end{bmatrix}.$$

Solution: We first look for eigenvalues, like always. We get these by computing

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} 1-\lambda & 9 & 6 \\ 0 & 1-\lambda & 6 \\ 0 & -3 & -5-\lambda \end{bmatrix}\right).$$

We will compute this by cofactor expansion along the second row.

$$\begin{aligned} \det\left(\begin{bmatrix} 1-\lambda & 9 & 6 \\ 0 & 1-\lambda & 6 \\ 0 & -3 & -5-\lambda \end{bmatrix}\right) &= (-1)^{2+2}(1-\lambda)\det\left(\begin{bmatrix} 1-\lambda & 6 \\ 0 & -5-\lambda \end{bmatrix}\right) \\ &\quad + (-1)^{2+3} 6 \det\left(\begin{bmatrix} 1-\lambda & 9 \\ 0 & -3 \end{bmatrix}\right) \\ &= (1-\lambda)(1-\lambda)(-5-\lambda) - 6(1-\lambda)(-3) \\ &= (1-\lambda)((1-\lambda)(-5-\lambda) + 18) \\ &= (1-\lambda)(\lambda^2 + 4\lambda + 13) \end{aligned}$$

so that one eigenvalue is at $\lambda = 1$. For the other two, we use the quadratic formula to obtain

$$\lambda = \frac{-4 \pm \sqrt{16 - 4 \cdot 13}}{2} = \frac{-4 \pm \sqrt{-36}}{2} = -2 \pm 3i.$$

Thus, we have one real eigenvalue and two complex eigenvalues.

For $\lambda = 1$, we know that the eigenvector must satisfy

$$\begin{bmatrix} 0 & 9 & 6 \\ 0 & 0 & 6 \\ 0 & -3 & -6 \end{bmatrix} \vec{v} = \vec{0}.$$

Row reduction will reduce this matrix to

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

(*Check this!*) so that the eigenvector in this case is

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

For the eigenvalue $-2 + 3i$, we get that the eigenvector must satisfy

$$\begin{bmatrix} 3-3i & 9 & 6 \\ 0 & 3-3i & 6 \\ 0 & -3 & -3-3i \end{bmatrix} \vec{v} = \vec{0}.$$

We now want to row reduce the coefficient matrix. To do so, we start by dividing the first row by 3 then multiplying by $1 + i$.

$$\begin{aligned} \begin{bmatrix} 3-3i & 9 & 6 \\ 0 & 3-3i & 6 \\ 0 & -3 & -3-3i \end{bmatrix} &\to \begin{bmatrix} 1-i & 3 & 2 \\ 0 & 3-3i & 6 \\ 0 & -3 & -3-3i \end{bmatrix} \\ &\to \begin{bmatrix} (1-i)(1+i) & 3(1+i) & 2(1+i) \\ 0 & 3-3i & 6 \\ 0 & -3 & -3-3i \end{bmatrix}. \\ &\to \begin{bmatrix} 2 & 3+3i & 2+2i \\ 0 & 3-3i & 6 \\ 0 & -3 & -3-3i \end{bmatrix} \end{aligned}$$

We could divide the first row by 2 to get to a 1 in the top-right entry, but we'll wait on that in order to avoid fractions. To row reduce the rest of the matrix, we will divide each of the remaining two rows by 3, and then multiply the second by $1 + i$, just like we did to the first row.

$$\begin{aligned} \begin{bmatrix} 2 & 3+3i & 2+2i \\ 0 & 3-3i & 6 \\ 0 & -3 & -3-3i \end{bmatrix} &\to \begin{bmatrix} 2 & 3+3i & 2+2i \\ 0 & 1-i & 2 \\ 0 & -1 & -1-i \end{bmatrix} \\ &\to \begin{bmatrix} 2 & 3+3i & 2+2i \\ 0 & 2 & 2+2i \\ 0 & -1 & -1-i \end{bmatrix} \\ &\to \begin{bmatrix} 2 & 3+3i & 2+2i \\ 0 & 1 & 1+i \\ 0 & -1 & -1-i \end{bmatrix} \end{aligned}$$

which illlustrates that the last two rows are redundant. Thus, the reduced form of the matrix that we have (which is not quite a row echelon form, but it is enough to back-solve for the eigenvector) is

$$\begin{bmatrix} 2 & 3+3i & 2+2i \\ 0 & 1 & 1+i \\ 0 & 0 & 0 \end{bmatrix}.$$

This means that the eigenvector $\vec{v}$ must satisfy

$$2v_1 + (3+3i)v_2 + (2+2i)v_3 = 0 \qquad v_2 + (1+i)v_3 = 0.$$

We can satisfy the second of these equations by choosing $v_2 = 1+i$ and $v_3 = -1$. Plugging these values into the first equation gives that

$$\begin{aligned} 0 &= 2v_1 + (3+3i)v_2 + (2+2i)v_3 \\ &= 2v_1 + (3+3i)(1+i) + (2+2i)(-1) \\ &= 2v_1 + 3 + 3i + 3i - 3 - 2 - 2i \\ &= 2v_1 - 2 + 4i \end{aligned}$$

Therefore, we need to take $v_1 = 1 - 2i$, giving that the eigenvector is

$$\begin{bmatrix} 1-2i \\ 1+i \\ -1 \end{bmatrix}.$$

A very similar computation following the same set of steps (or just using the remark below) for the eigenvalue $-2-3i$ gives that this corresponding eigenvector is

$$\begin{bmatrix} 1+2i \\ 1-i \\ -1 \end{bmatrix}.$$

One fact that comes out of those examples is that the eigenvectors for conjugate eigenvalues are also complex conjugates. This comes from the fact that A is a real matrix, which means that if

$$A\vec{v} = \lambda\vec{v}$$

and we take the complex conjugate of both sides, we get that

$$A\bar{\vec{v}} = \bar{A}\bar{\vec{v}} = \overline{\lambda\vec{v}} = \bar{\lambda}\bar{\vec{v}}$$

so that $\bar{\vec{v}}$ is an eigenvector for $\bar{\lambda}$. This means that, when solving these types of problems, we only need to find one of the complex eigenvectors and can get the other by taking the complex conjugate.

Repeated Eigenvalues

Distinct and complex eigenvalues all work out nicely and in pretty much the same manner. For repeated eigenvalues, the issues get more significant.

Example 86. *Find the eigenvalues and eigenvectors of the matrices*

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

Solution: For the matrix A, we can compute the characteristic polynomial

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

Therefore, we have a double root at 3 for this matrix. Therefore, the only eigenvalue we get is 3. When we look to find the eigenvectors, we get

$$A - 3I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

so that this matrix multiplied by *any* vector is zero. Therefore, we can use both $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ as eigenvectors.

On the other hand, the matrix B has a characteristic polynomial

$$\begin{aligned}\det(B-\lambda I) &= \det\left(\begin{bmatrix}4-\lambda & -1\\ 1 & 2-\lambda\end{bmatrix}\right)\\ &= (4-\lambda)(2-\lambda)-(-1)(1) = \lambda^2-6\lambda+8+1\\ &= \lambda^2-6\lambda+9 = (\lambda-3)^2\end{aligned}$$

so again, we have a double root at 3. However, when we go to find the eigenvectors, we get that

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

which gives that an eigenvector must satisfy $v_1 - v_2 = 0$ so $\begin{bmatrix}1\\1\end{bmatrix}$ works.

There is a big difference between these two examples. Both had the same characteristic polynomial of $(\lambda-3)^2$, but for A, we could find two linearly independent eigenvectors, but for B, we could only find 1. This seems like it might be a problem, since we would like to get to two eigenvectors like we did for both of the previous two cases. This leads us to define the following for A and $n \times n$ matrix and r an eigenvalue of A.

Definition 24.

- *The* algebraic multiplicity *of r is the power of $(\lambda - r)$ in the characteristic polynomial if A.*
- *The* geometric multiplicity *of r is the number of linearly independent eigenvectors of A with eigenvalue r.*
- *The* defect *of r is the difference between the algebraic multiplicity and the geometric multiplicity of r.*
- *We say that an eigenvalue is* defective *if the defect is at least 1.*

For the previous example, the algebraic multiplicity of 3 for both A and B was 2, but the geometric multiplictiy of 3 for A is 2, and for B is it only 1. Therefore A has a defect of 0 and B has a defect of 1, so 3 is a defective eigenvalue for matrix B.

In terms of these multiplicities, there are two facts that are known to be true.

(a) If r is an eigenvalue, then both the algebraic and geometric multiplicity are at least 1.

(b) The algebraic multiplicity of any eigenvalue is always greater than or equal to the geometric multiplicity.

This tells us that in the case of real and distinct eigenvalues, every eigenvalue has multiplicity 1. Since the geometric multiplicity is also 1, this means that none of these eigenvalues are defective. This was great, because it let us get to n eigenvectors for an $n \times n$ matrix, and these generated a basis of $\mathbb{R}^n$.

Why is a defective eigenvalue a problem? When we go solve differential equations using the method in Chapter **??**, having a 'full set' of eigenvectors, or n eigenvectors for an $n \times n$ matrix, will be very important. When we have a defective eigenvalue, we can't get there. Since the degree of the characteristic polynomial is n, the only way we get to n eigenvectors is if every eigenvalue has a number of linearly independent eigenvectors equal to the algebraic multiplicity, which means they are not defective.

So how can we fix this? Well, there's not really much we can do in the way of finding more eigenvectors, because they don't exist. The replacement that we have is, in linear algebra contexts, called a *generalized eigenvector*. We will see this idea come back up in § **??** in a more natural way. The rest of this section contains a more detailed definition of generalized eigenvectors. You are welcome to skip this part on a first reading and come back after you are more comfortable with eigenvalues and eigenvectors, or when the material comes back around again in § **??**.

If r is an defective eigenvalue of the matrix A with eigenvector $\vec{v}$, a *generalized eigenvector* of A is a vector $\vec{w}$ so that $(A-rI)\vec{w} = \vec{v}$. This is the same as the normal eigenvector equation with $\vec{v}$ on the right-hand side instead of $\vec{0}$. Since $(A-rI)\vec{v} = \vec{0}$, this also means that

$$(A-rI)^2\vec{w} = 0.$$

More generally, a generalized eigenvector is a vector $\vec{w}$ where there is a power $k \geq 1$ so that

$$(A-rI)^k\vec{w} = 0 \qquad \text{but} \qquad (A-rI)^{k-1}\vec{w} \neq 0.$$

It might seems strange where this comes from, but we will see why this formula makes more sense once we try to solve differential equations using matrices in § **??**.

Example 87. *Find a generalized eigenvector of eigenvalue* 3 *for the matrix*

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

Solution: Previously, we found that $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector for B with eigenvalue 3. To find a generalized eigenvector, we need a vector $\vec{w}$ so that

$$(B - 3I)\vec{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Plugging in the matrix for $B - 3I$ gives that we need

$$\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Both of the rows of this matrix becomes the equation

$$w_1 - w_2 = 1.$$

There are many values of w_1 and w_2 that make this work. We can pick $w_1 = 1$ and $w_2 = 0$. This will give a generalized eigenvector of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. We could also pick $w_1 = 3$ and $w_2 = 2$, to get a generalized eigenvector as $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$. Any of these choices work as a generalized eigenvector.

Example 88. *Find the eigenvalues and eigenvectors (and generalized eigenvectors if needed) of the matrix*

$$A = \begin{bmatrix} -2 & 0 & 1 \\ 19 & 2 & -16 \\ -1 & 0 & 0 \end{bmatrix}.$$

Solution: We start by looking for the eigenvalues through the characteristic polynomial.

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} -2-\lambda & 0 & 1 \\ 19 & 2-\lambda & -16 \\ -1 & 0 & 0-\lambda \end{bmatrix} \right)$$

To compute this determinant, we will expand along column 2, because it only has one non-zero entry. This gives

$$\begin{aligned} \det(A - \lambda I) &= (-1)^{2+2}(2-\lambda)\det\left(\begin{bmatrix} -2-\lambda & 1 \\ -1 & -\lambda \end{bmatrix} \right) \\ &= (2-\lambda)((-2-\lambda)(-\lambda) + 1) \\ &= (2-\lambda)(\lambda^2 + 2\lambda + 1) = (2-\lambda)(\lambda+1)^2 \end{aligned}$$

so we have an eigenvalue at 2 and a double eigenvalue at -1.

First, let's look for the eigenvector for eigenvalue 2. In this case, we know that the eigenvector must satisfy

$$\begin{bmatrix} -4 & 0 & 1 \\ 19 & 0 & -16 \\ -1 & 0 & -2 \end{bmatrix} \vec{v} = \vec{0}.$$

Row reducing the coefficient matrix will give

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

so that a corresponding eigenvector is

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

since we know that $v_1 = 0$ and $v_3 = 0$.

For $\lambda = -1$, we see that an eigenvector must satisfy

$$\begin{bmatrix} -1 & 0 & 1 \\ 19 & 3 & -16 \\ -1 & 0 & 1 \end{bmatrix} \vec{v} = \vec{0}.$$

We now look to row reduce this coefficient matrix.

$$\begin{aligned} \begin{bmatrix} -1 & 0 & 1 \\ 19 & 3 & -16 \\ -1 & 0 & 1 \end{bmatrix} &\to \begin{bmatrix} 1 & 0 & -1 \\ 19 & 3 & -16 \\ -1 & 0 & 1 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 0 \end{bmatrix} \quad . \\ &\to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Therefore, we know that

$$v_1 - v_3 = 0 \qquad v_2 + v_3 = 0.$$

If we pick $v_3 = 1$, then we know that $v_2 = -1$ and $v_1 = 1$, so the only eigenvector we get for $\lambda = -1$ is

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

Since we only found one eigenvector for $\lambda = -1$ and $\lambda + 1$ was squared in the characteristic polynomial, this is a defective eigenvalue. Thus, we can look for a generalized eigenvalue here, which means that we need to solve for a vector $\vec{w}$ with

$$\begin{bmatrix} -1 & 0 & 1 \\ 19 & 3 & -16 \\ -1 & 0 & 1 \end{bmatrix} \vec{w} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

We can then row reduce the augmented matrix to see what we can pick for $\vec{w}$.

$$\begin{aligned} \begin{bmatrix} -1 & 0 & 1 & 1 \\ 19 & 3 & -16 & -1 \\ -1 & 0 & 1 & 1 \end{bmatrix} &\to \begin{bmatrix} 1 & 0 & -1 & -1 \\ 19 & 3 & -16 & -1 \\ -1 & 0 & 1 & 1 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & 0 & -1 & -1 \\ 0 & 3 & 3 & 18 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &\to \begin{bmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Thus, the generalized eigenvector $\vec{w}$ must satisfy

$$w_1 - w_3 = -1 \qquad w_2 + w_3 = 6.$$

We can pick any non-zero numbers to do this, so we can take $w_3 = 1$, $w_2 = 5$ and $w_1 = 0$. Thus, the generalized eigenvector here is

$$\begin{bmatrix} 0 \\ 5 \\ 1 \end{bmatrix}.$$

linearAlgebra/practice/eigenvalues-practice1.tex

Practice for Eigenvalues

Why?

Exercise 473 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} -8 & -18 \\ 4 & 10 \end{bmatrix}$$

Exercise 474 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} -2 & 0 \\ 8 & -4 \end{bmatrix}$$

Exercise 475 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} -7 & 1 \\ -12 & 0 \end{bmatrix}$$

Exercise 476 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} -3 & 5 \\ -8 & 9 \end{bmatrix}$$

Exercise 477 *Find the eigenvalues and corresponding eigenvectors of the matrix*

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

Exercise 478 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} -4 & 1 \\ -8 & 0 \end{bmatrix}$$

Author(s): Matthew Charnley and Jason Nowell

Exercise 479 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} 0 & -8 \\ 2 & 8 \end{bmatrix}$$

Exercise 480 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} 1 & -2 \\ 8 & -7 \end{bmatrix}$$

Exercise 481 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} 4 & 0 & 0 \\ -4 & 2 & 1 \\ -6 & 0 & 1 \end{bmatrix}$$

Exercise 482 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} -4 & 9 & 9 \\ -3 & 6 & 9 \\ 3 & -7 & -10 \end{bmatrix}$$

Exercise 483 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} -2 & 0 & 0 \\ 0 & 4 & 6 \\ 6 & -3 & -2 \end{bmatrix}$$

Exercise 484 *Find the eigenvalues and corresponding eigenvectors of the matrix*

$$\begin{bmatrix} 5 & 3 & 6 \\ 2 & 2 & 2 \\ -3 & -2 & -3 \end{bmatrix}$$

Exercise 485 *Find the eigenvalues and eigenvectors for the matrix below. Compute generalized eigenvectors if needed to get to a total of two vectors.*

$$\begin{bmatrix} -11 & -9 \\ 12 & 10 \end{bmatrix}$$

Exercise 486 *Find the eigenvalues and eigenvectors for the matrix below. Compute generalized eigenvectors if needed to get to a total of two vectors.*

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

Exercise 487 *This exercise will work through the process of finding the eigenvalues and corresponding eigenvectors of the matrix*

$$A = \begin{bmatrix} -2 & 0 & -3 \\ 12 & 5 & 12 \\ 0 & -1 & 1 \end{bmatrix}.$$

a) *Find the characteristic polynomial of this matrix by computing* $\det(A - \lambda I)$ *using any method from this section.*

b) *This polynomial can be rewritten as* $-(\lambda - r_1)^2(\lambda - r_2)$ *where* r_1 *and* r_2 *are the eigenvalues of* A*. What are the eigenvalues? What is each of their algebraic multiplicity? (Hint: One of these roots is 2)*

c) *Find an eigenvector for eigenvalue* r_2 *above. What is the geometric multiplicity of this eigenvalue?*

d) *Find an eigenvector for eigenvalue* r_1*. What is the geometric multiplicity of this eigenvalue?*

e) *There is only one possible eigenvector for* r_1*, which means it is defective. Find a solution to the equation* $(A - r_1 I)\vec{w} = \vec{v}$*, where* $\vec{v}$ *is the eigenvector you found in the previous part. This is the generalized eigenvector for* r_1*.*

Exercise 488 *We say that a matrix* A *is* diagonalizable *if there exist matrices* D *and* P *so that* $PDP^{-1} = A$*. This really means that* A *can be represented by a diagonal matrix in a different basis (as opposed to the standard basis). One way this can be done is with eigenvalues.*

a) *Consider the matrix* A *given by*

$$A = \begin{bmatrix} -4 & 6 \\ -1 & 1 \end{bmatrix}.$$

Find the eigenvalues and corresponding eigenvectors of this matrix.

b) *Form two matrices,* D*, a diagonal matrix with the eigenvalues of* A *on the diagonal, and* E*, a matrix whose columns are the eigenvectors of* A *in the* same order *as the eigenvalues were put into* D*. Write out these matrices.*

c) *Compute* E^{-1}*.*

d) *Work out the products* EDE^{-1} and $E^{-1}AE$*. What do you notice here?*

This shows that, in the case of a 2×2 *matrix, if we have two distinct real eigenvalues, that matrix is diagonalizable, using the eigenvectors.*

Exercise 489 *Follow the process outlined in Exercise to attempt to diagonalize the matrix*

$$\begin{bmatrix} 13 & 14 & 12 \\ -6 & -4 & -6 \\ -3 & -6 & -2 \end{bmatrix}$$

Hint: 1 is an eigenvalue.

Exercise 490 *The diagonalization process decribed in Exercise works for any case where there are real and distinct eigenvalues, as well as complex eigenvalues (but the algebra with the complex numbers gets complicated). It may or may not work in the case of repeated eigenvalues, and it fails whenever there are defective eigenvalues. Consider the matrix*

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

a) *Find the eigenvalue(s) of this matrix, and see that we have a repeated eigenvalue.*

b) *Find the eigenvector for that eigenvalue, as well as a generalized eigenvector.*

c) *Build a matrix E like before, but this time put the eigenvector in the first column and the generalized eigenvector in the second. Compute E^{-1}.*

d) *Find the product $E^{-1}AE$. Before, this gave us a diagonal matrix, but what do we get now?*

The matrix we get here is almost diagonal, but not quite. It turns out that this is the best we can do for matrices with defective eigenvalues. This matrix is often called J and is the Jordan Form *of the matrix A.*

Exercise 491 *Follow the process in Exercise to find the Jordan Form of the matrix*

$$\begin{bmatrix} -7 & 5 & 5 \\ -4 & 5 & 7 \\ -6 & 3 & 1 \end{bmatrix}.$$

Related Topics in Linear Algebra

We discuss Related Topics in Linear Algebra

Subspaces and span

Assume that we find two vectors that solve $A\vec{x} = 0$. What other vectors also solve this equation? In our discussion of linear combinations, we saw that if $\vec{x}_1$ and $\vec{x}_2$ solve $A\vec{x} = 0$, then so does $A(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2)$ for any constants α_1 and α_2. Thus, all linear combinations will also solve the equation. This leads to the definition of the span of a set of vectors.

Definition 25. *The set of all linear combinations of a set of vectors is called their* span*.*

$$\operatorname{span}\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\} = \{\textit{Set of all linear combinations of } \vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}.$$

Thus, if two vectors solve a homogeneous equation, so does everything in the span of those two vectors. The span of a collection of vectors is an example of a subspace, which is a common object in linear algebra. We say that a set S of vectors in $\mathbb{R}^n$ is a *subspace* if whenever $\vec{x}$ and $\vec{y}$ are members of S and α is a scalar, then

$$\vec{x} + \vec{y}, \qquad \text{and} \qquad \alpha\vec{x}$$

are also members of S. That is, we can add and multiply by scalars and we still land in S. So every linear combination of vectors of S is still in S. That is really what a subspace is. It is a subset where we can take linear combinations and still end up being in the subset.

Example 89. *If we let $S = \mathbb{R}^n$, then this S is a subspace of $\mathbb{R}^n$. Adding any two vectors in $\mathbb{R}^n$ gets a vector in $\mathbb{R}^n$, and so does multiplying by scalars.*

The set $S' = \{\vec{0}\}$, that is, the set of the zero vector by itself, is also a subspace of $\mathbb{R}^n$. There is only one vector in this subspace, so we only need to check for that one vector, and everything checks out: $\vec{0} + \vec{0} = \vec{0}$ and $\alpha\vec{0} = \vec{0}$.

The set S'' of all the vectors of the form (a, a) for any real number a, such as $(1, 1)$, $(3, 3)$, or $(-0.5, -0.5)$ is a subspace of $\mathbb{R}^2$. Adding two such vectors, say $(1, 1) + (3, 3) = (4, 4)$ again gets a vector of the same form, and so does multiplying by a scalar, say $8(1, 1) = (8, 8)$.

We can apply these ideas to the vectors that live inside a matrix. The span of the rows of a matrix A is called the *row space*. The row space of A and the row space of the row echelon form of A are the same, because reducing the matrix A to its row echelon form involves taking linear combinations, which will preserve the span. In the example,

$$\begin{aligned} \text{row space of } \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} &= \operatorname{span}\left\{ \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}, \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}, \begin{bmatrix} 7 & 8 & 9 \end{bmatrix} \right\} \\ &= \operatorname{span}\left\{ \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 \end{bmatrix} \right\}. \end{aligned}$$

Similarly to row space, the span of columns is called the *column space*.

$$\text{column space of } \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} \right\}.$$

In particular, to find a set of linearly independent columns we need to look at where the pivots were. If you recall above, when solving $A\vec{x} = \vec{0}$ the key was finding the pivots, any non-pivot columns corresponded to free variables. That means we can solve for the non-pivot columns in terms of the pivot columns. Let's see an example.

Learning outcomes: Compute the rank of a matrix Find a maximal linearly independent subset of a set of vectors Compute a basis of a subspace and the dimension of that subspace Determine the kernel of a matrix using row reduction Understand the connection between rank and nullity in a given matrix Compute the inverse of a matrix using row reduction Use properties of the trace and determinant to analyze the eigenvalues of a matrix.

Author(s): Matthew Charnley and Jason Nowell

Example 90. *Find the linearly independent columns of the matrix*

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

Solution: We find a pivot and reduce the rows below:

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

We find the next pivot, make it one, and rinse and repeat:

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

The final matrix is the row echelon form of the matrix. Consider the pivots that we marked. The pivot columns are the first and the third column. All other columns correspond to free variables when solving $A\vec{x} = \vec{0}$, so all other columns can be solved in terms of the first and the third column. In other words

$$\text{column space of } \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix} \right\} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}.$$

We could perhaps use another pair of columns to get the same span, but the first and the third are guaranteed to work because they are pivot columns.

In the previous example, this means that only the first and third colums are "important" in the sense of generating the full column space as a span. We would like to have a way to talk about what these first and third columns do.

Definition 26 (Spanning set)**.** *Let S be a subspace of a vector space. The set $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_k\}$ is a* spanning set *for the subspace S if each of these vectors are in S and the span of $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_k\}$ is equal to S.*

In the context of the previous example, for the matrix

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

we know that

$$\text{column space of } \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix} \right\} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}.$$

This means that both

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix} \right\} \quad \text{and} \quad \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}$$

are spanning sets for this column space.

The idea also works in reverse. Suppose we have a bunch of column vectors and we just need to find a linearly independent set. For example, suppose we started with the vectors

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix}.$$

These vectors are not linearly independent as we saw above. In particular, the span $\vec{v}_1$ and $\vec{v}_3$ is the same as the span of all four of the vectors. So $\vec{v}_2$ and $\vec{v}_4$ can both be written as linear combinations of $\vec{v}_1$ and $\vec{v}_3$. A common thing that comes up in practice is that one gets a set of vectors whose span is the set of solutions of some problem. But perhaps we get way too many vectors, we want to simplify. For example above, all vectors in the span of $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ can be written $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \alpha_3\vec{v}_3 + \alpha_4\vec{v}_4$ for some numbers $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. But it is also true that every such vector can be written as $a\vec{v}_1 + b\vec{v}_3$ for two numbers a and b. And one has to admit, that looks much simpler. Moreover, these numbers a and b are unique. More on that later in this section.

To find this linearly independent set we simply take our vectors and form the matrix $[\vec{v}_1\ \vec{v}_2\ \vec{v}_3\ \vec{v}_4]$, that is, the matrix

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

We crank up the row-reduction machine, feed this matrix into it, and find the pivot columns and pick those. In this case, $\vec{v}_1$ and $\vec{v}_3$.

Basis and dimension

At this point, we have talked about subspaces, and two other properties of sets of vectors: linear independence and being a spanning set for a subspace. In some sense, these two properties are in opposition to each other. If I add more vectors to a set, I am more likely to become a spanning set (because I have more options for adding to get other vectors), but less likely to be independent (because there are more possibilities for a linear combination to be zero). Similarly, the reverse is true; removing vectors means the set is more likely to be linearly independent, but less likely to span a given subspace. The question then becomes if there is a sweet spot where both things are true, and that leads to the definition of a basis.

Definition 27. *If S is a subspace and we can find k linearly independent vectors in S*

$$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k,$$

such that every other vector in S is a linear combination of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$, then the set $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is called a basis *of S. In other words, S is the span of $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$. We say that S has* dimension *k, and we write*

$$\dim S = k.$$

The next theorem illustrates the main properties and classification of a basis of a vector space.

Theorem 22. *If $S \subset \mathbb{R}^n$ is a subspace and S is not the trivial subspace $\{\vec{0}\}$, then there exists a unique positive integer k (the dimension) and a (not unique) basis $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$, such that every $\vec{w}$ in S can be uniquely represented by*

$$\vec{w} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_k\vec{v}_k,$$

for some scalars α_1, α_2, ..., α_k.

We should reiterate that while k is unique (a subspace cannot have two different dimensions), the set of basis vectors is not at all unique. There are lots of different bases for any given subspace. Finding just the right basis for a subspace is a large part of what one does in linear algebra. In fact, that is what we spend a lot of time on in linear differential equations, although at first glance it may not seem like that is what we are doing.

Example 91. *The standard basis*

$$\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n,$$

is a basis of $\mathbb{R}^n$ (hence the name). So as expected

$$\dim \mathbb{R}^n = n.$$

On the other hand the subspace $\{\vec{0}\}$ is of dimension 0.

The subspace S'' from a previous example, that is, the set of vectors (a, a) is of dimension 1. One possible basis is simply $\{(1, 1)\}$, the single vector $(1, 1)$: every vector in S'' can be represented by $a(1, 1) = (a, a)$. Similarly another possible basis would be $\{(-1, -1)\}$. Then the vector (a, a) would be represented as $(-a)(-1, -1)$. In this case, the subspace S'' has many different bases, two of which are $\{(1, 1)\}$ and $\{(-1, -1)\}$, and the vector (a, a) has a different representation (different constant) for the different bases.

Row and column spaces of a matrix are also examples of subspaces, as they are given as the span of vectors. We can use what we know about row spaces and column spaces from the previous section to find a basis.

Example 92. *Earlier, we considered the matrix*

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

Using row reduction to find the pivot columns, we found

$$\textit{column space of } A \left(\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 6 & 7 & 8 \end{bmatrix} \right) = \operatorname{span} \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix} \right\}.$$

What we did was we found the basis of the column space. The basis has two elements, and so the column space of A is two dimensional.

We would have followed the same procedure if we wanted to find the basis of the subspace X spanned by

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix}.$$

We would have simply formed the matrix A with these vectors as columns and repeated the computation above. The subspace X is then the column space of A.

Example 93. *Consider the matrix*

$$L = \begin{bmatrix} 1 & 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix}$$

Conveniently, the matrix is in reduced row echelon form. The column space is the span of the pivot columns, because the pivot columns always form a basis for the column space of a matrix. It is the 3-dimensional space

$$\textit{column space of } L = \operatorname{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^3.$$

The row space is the 3-dimensional space

$$\textit{row space of } L = \operatorname{span} \left\{ \begin{bmatrix} 1 & 2 & 0 & 0 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 & 5 \end{bmatrix} \right\}.$$

As these vectors have 5 components, we think of the row space of L as a subspace of $\mathbb{R}^5$.

Rank

In that last example, we noticed that the dimension of the row space and the column space were the same. It turns out that this is not a coincidence. In order to describe this in more detail, we need to define one more term.

Definition 28. *Given a matrix A, the maximal number of linearly independent rows is called the* rank *of A, and we write "*rank A*" for the rank.*

For example,

$$\operatorname{rank} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix} = 1.$$

The second and third row are multiples of the first one. We cannot choose more than one row and still have a linearly independent set. But what is

$$\operatorname{rank} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad = \quad ?$$

That seems to be a tougher question to answer. The first two rows are linearly independent, so the rank is at least two. If we would set up the equations for the α_1, α_2, and α_3, we would find a system with infinitely many solutions. One solution is

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

So the set of all three rows is linearly dependent, the rank cannot be 3. Therefore the rank is 2.

But how can we do this in a more systematic way? We find the row echelon form!

$$\text{Row echelon form of} \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad \text{is} \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

The elementary row operations do not change the set of linear combinations of the rows (that was one of the main reasons for defining them as they were). In other words, the span of the rows of the A is the same as the span of the rows of the row echelon form of A. In particular, the number of linearly independent rows is the same. And in the row echelon form, all nonzero rows are linearly independent. This is not hard to see. Consider the two nonzero rows in the example above. Suppose we tried to solve for the α_1 and α_2 in

$$\alpha_1 \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}.$$

Since the first column of the row echelon matrix has zeros except in the first row means that $\alpha_1 = 0$. For the same reason, α_2 is zero. We only have two nonzero rows, and they are linearly independent, so the rank of the matrix is 2. This also tells us that if we were trying to solve the system of equations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= a \\ 4x_1 + 5x_2 + 6x_3 &= b \\ 7x_1 + 8x_2 + 9x_3 &= c \end{aligned}$$

we would get that one row of the reduced augmented matrix has all zeros on the left side, and so this system either has a free variable or is inconsistent, because only two equations here are relevant.

Referring back to the examples from earlier in this section, we could carry out the same calculations to say that

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

and

$$\operatorname{rank} \begin{bmatrix} 1 & 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix} = 3.$$

We know how to find the set of linearly independent rows, but sometimes it may also be useful to find the linearly independent columns as well. It is a tremendously useful fact that the number of linearly independent columns is always the same as the number of linearly independent rows:

Theorem 23. $\operatorname{rank} A = \operatorname{rank} A^T$

Or, in the context of the row and column spaces that we have already discussed:

Theorem 24 (Rank)**.** *The dimension of the column space and the dimension of the row space of a matrix A are both equal to the rank of A.*

This relates to the statement at the start of this section; since the number of vectors that we needed to take to get a basis of linearly independent columns was always the same as the number of pivots, and the number of pivots is the rank, we get that above theorem.

Kernel

The set of solutions of a linear equation $L\vec{x} = \vec{0}$, the kernel of L, is a subspace: If $\vec{x}$ and $\vec{y}$ are solutions, then

$$L(\vec{x} + \vec{y}) = L\vec{x} + L\vec{y} = \vec{0} + \vec{0} = \vec{0}, \qquad \text{and} \qquad L(\alpha\vec{x}) = \alpha L\vec{x} = \alpha\vec{0} = \vec{0}.$$

So $\vec{x} + \vec{y}$ and $\alpha\vec{x}$ are solutions. The dimension of the kernel is called the *nullity* of the matrix.

The same sort of idea governs the solutions of linear differential equations. We try to describe the kernel of a linear differential operator, and as it is a subspace, we look for a basis of this kernel. Much of this book is dedicated to finding such bases.

The kernel of a matrix is the same as the kernel of its reduced row echelon form. For a matrix in reduced row echelon form, the kernel is rather easy to find. If a vector $\vec{x}$ is applied to a matrix L, then each entry in $\vec{x}$ corresponds to a column of L, the column that the entry multiplies. To find the kernel, pick a non-pivot column make a vector that has a -1 in the entry corresponding to this non-pivot column and zeros at all the other entries corresponding to the other non-pivot columns. Then for all the entries corresponding to pivot columns make it precisely the value in the corresponding row of the non-pivot column to make the vector be a solution to $L\vec{x} = \vec{0}$. This procedure is best understood by example.

Example 94. *Consider*

$$L = \begin{bmatrix} 1 & 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix}.$$

This matrix is in reduced row echelon form, the pivots are marked. There are two non-pivot columns, so the kernel has dimension 2, that is, it is the span of 2 vectors. Let us find the first vector. We look at the first non-pivot column, the 2^{nd} column, and we put a -1 in the 2^{nd} entry of our vector. We put a 0 in the 5^{th} entry as the 5^{th} column is also a non-pivot column:

$$\begin{bmatrix} ? \\ -1 \\ ? \\ ? \\ 0 \end{bmatrix}.$$

Let us fill the rest. When this vector hits the first row, we get a -2 and 1 times whatever the first question mark is. So make the first question mark 2. For the second and third rows, it is sufficient to make it the question marks zero. We are really filling in the non-pivot column into the remaining entries. Let us check while marking which numbers went where:

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

Yay! How about the second vector. We start with

$$\begin{bmatrix} ? \\ 0 \\ ? \\ ? \\ -1. \end{bmatrix}$$

We set the first question mark to 3, the second to 4, and the third to 5. Let us check, marking things as previously,

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

There are two non-pivot columns, so we only need two vectors. We have found the basis of the kernel. So,

$$\textit{kernel of } L = \operatorname{span} \left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 4 \\ 5 \\ -1 \end{bmatrix} \right\}$$

What we did in finding a basis of the kernel is we expressed all solutions of $L\vec{x} = \vec{0}$ as a linear combination of some given vectors.

The procedure to find the basis of the kernel of a matrix L:

(a) Find the reduced row echelon form of L.

(b) Write down the basis of the kernel as above, one vector for each non-pivot column.

The rank of a matrix is the dimension of the column space, and that is the span of the pivot columns, while the kernel is the span of vectors in the non-pivot columns. So the two numbers must add to the number of columns.

Theorem 25 (Rank–Nullity)**.** *If a matrix A has n columns, rank r, and nullity k (dimension of the kernel), then*

$$n = r + k.$$

The theorem is immensely useful in applications. It allows one to compute the rank r if one knows the nullity k and vice versa, without doing any extra work.

Let us consider an example application, a simple version of the so-called *Fredholm alternative.* A similar result is true for differential equations. Consider

$$A\vec{x} = \vec{b},$$

where A is a square $n \times n$ matrix. There are then two mutually exclusive possibilities:

(a) A nonzero solution $\vec{x}$ to $A\vec{x} = \vec{0}$ exists.

(b) The equation $A\vec{x} = \vec{b}$ has a unique solution $\vec{x}$ for every $\vec{b}$.

How does the Rank–Nullity theorem come into the picture? Well, if A has a nonzero solution $\vec{x}$ to $A\vec{x} = \vec{0}$, then the nullity k is positive. But then the rank $r = n - k$ must be less than n. In particular it means that the column space of A is of dimension less than n, so it is a subspace that does not include everything in $\mathbb{R}^n$. So $\mathbb{R}^n$ has to contain some vector $\vec{b}$ not in the column space of A. In fact, most vectors in $\mathbb{R}^n$ are not in the column space of A.

The idea of a kernel also comes up when defining and discussing eigenvectors. In order to find this vector, we are looking for a vector $\vec{v}$ so that

$$(A - \lambda I)\vec{v} = \vec{0}.$$

This means that we are looking for a vector $\vec{v}$ that is in the kernel of the matrix $(A - \lambda I)$. Since the kernel is also a subspace, this means that the set of all eigenvectors of a matrix A with a certain eigenvalue is a subspace, so it has a dimension. This dimension is number of linearly independent eigenvectors with that eigenvalue, so it is the geometric multiplicity of this eigenvalue. This also motivates why this is sometimes called the *eigenspace* for a given eigenvalue. Finding a basis of this subspace (which is also finding the kernel of the matrix $A - \lambda I$) is the exact same as the process of finding the eigenvectors of the matrix A.

Computing the inverse

If the matrix A is square and there exists a unique solution $\vec{x}$ to $A\vec{x} = \vec{b}$ for any $\vec{b}$ (there are no free variables), then A is invertible.

In particular, if $A\vec{x} = \vec{b}$ then $\vec{x} = A^{-1}\vec{b}$. Now we just need to compute what A^{-1} is. We can surely do elimination every time we want to find $A^{-1}\vec{b}$, but that would be ridiculous. The mapping A^{-1} is linear and hence given by a matrix, and we have seen that to figure out the matrix we just need to find where does A^{-1} take the standard basis vectors $\vec{e}_1$, $\vec{e}_2$, $\ldots$, $\vec{e}_n$.

That is, to find the first column of A^{-1} we solve $A\vec{x} = \vec{e}_1$, because then $A^{-1}\vec{e}_1 = \vec{x}$. To find the second column of A^{-1} we solve $A\vec{x} = \vec{e}_2$. And so on. It is really just n eliminations that we need to do. But it gets even easier. If you think about it, the elimination is the same for everything on the left side of the augmented matrix. Doing n eliminations separately we would redo most of the computations. Best is to do all at once.

Therefore, to find the inverse of A, we write an $n \times 2n$ augmented matrix $[\,A \mid I\,]$, where I is the identity matrix, whose columns are precisely the standard basis vectors. We then perform row reduction until we arrive at the reduced row echelon form. If A is invertible, then pivots can be found in every column of A, and so the reduced row echelon form of $[\,A \mid I\,]$ looks like $[\,I \mid A^{-1}\,]$. We then just read off the inverse A^{-1}. If you do not find a pivot in every one of the first n columns of the augmented matrix, then A is not invertible.

This is best seen by example.

Example 95. *Find the inverse of the matrix*

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

Solution: We write the augmented matrix and we start reducing:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -4 & -5 & -2 & 1 & 0 \\ 0 & -5 & -9 & -3 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 5/4 & 1/2 & 1/4 & 0 \\ 0 & -5 & -9 & -3 & 0 & 1 \end{array}\right] \to$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 5/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & -11/4 & -1/2 & -5/4 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 5/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 1 & 2/11 & 5/11 & -4/11 \end{array}\right] \to$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 5/11 & -5/11 & 12/11 \\ 0 & 1 & 0 & 3/11 & -9/11 & 5/11 \\ 0 & 0 & 1 & 2/11 & 5/11 & -4/11 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/11 & 3/11 & 2/11 \\ 0 & 1 & 0 & 3/11 & -9/11 & 5/11 \\ 0 & 0 & 1 & 2/11 & 5/11 & -4/11 \end{array}\right].$$

So

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} -1/11 & 3/11 & 2/11 \\ 3/11 & -9/11 & 5/11 \\ 2/11 & 5/11 & -4/11 \end{bmatrix}.$$

$\lrcorner$

Not too terrible, no? Perhaps harder than inverting a 2×2 matrix for which we had a formula, but not too bad. Really in practice this is done efficiently by a computer.

Trace and Determinant of Matrices

The next thing to add into our toolbox of matrices is the idea of the trace of a matrix, and how it and the determinant relate to the eigenvalues of said matrix.

Definition 29. *Let A be an $n \times n$ square matrix. The* trace *of A is the sum of all diagonal entries of A.*

For example, if we have the matrix

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

the trace is $1 + 2 + 3 = 6$.

The trace is important in our context because it also tells us something about the eigenvalues of a matrix. To work this out, let's consider the generic 2×2 matrix and how we would find the eigenvalues. If we have a 2×2 matrix of the form

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

we can write out the expression $\det(A - \lambda I)$ in order to find the eigenvalues. In this case, we would get

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix}\right) = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc).$$

However, the coefficients in this polynomial look familiar. $(ad - bc)$ is just the determinant of the matrix A, and $a + d$ is the trace. Therefore, for any 2×2 matrix, we could write the characteristic polynomial as

$$\det(A - \lambda I) = \lambda^2 - T\lambda + D \tag{31}$$

where T is the trace of the matrix and D is the determinant. On the other hand, assume that r_1 and r_2 are the two eigenvalues of this matrix (whether they be real, complex, or repeated). In that case, we know that this polynomial has r_1 and r_2 as roots. Therefore, it is equal to

$$\det(A - \lambda I) = (\lambda - r_1)(\lambda - r_2) = \lambda^2 - (r_1 + r_2)\lambda + r_1 r_2. \tag{32}$$

Matching up the coefficient of λ and the constant term in (31) and (32) gives the relation that

$$T = r_1 + r_2 \qquad D = r_1 r_2,$$

that is, the trace of the matrix is the sum of the eigenvalues, and the determinant of the matrix is the product of the eigenvalues. We only showed this fact for 2×2 matrices, but it does hold for matrices of all sizes, giving us the following theorem.

Theorem 26. *Let A be an $n \times n$ square matrix with eigenvalues $\lambda_1,\ \lambda_2,\ ..., \lambda_n$, written with multiplicity if needed. Then*

(a) *The trace of A is $\lambda_1 + \lambda_2 + \cdots + \lambda_n$.*

(b) *The determinant of A is $(\lambda_1)(\lambda_2)\cdots(\lambda_n)$.*

From the above statement, we note that if any of the eigenvalues is zero, the product of all eigenvalues will be zero, and so the matrix will have zero determinant. This gives an extra follow-up fact, and addition to Theorem 20.

Theorem 27. *A matrix A is invertible if and only if all of it's eigenvalues are non-zero.*

Example 96. *Use the facts above to analyze the eigenvalues of the matrix*

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}.$$

Solution: From the matrix A, we can compute that the trace of A is $1+4 = 5$, and the determinant is $(1)(4)-(2)(5) = -6$. Based on the theorem above, we know that the two eigenvalues of this matrix must add to 5 and multiply to -6. While you could probably guess the numbers here, the important take-aways from this example are what we can learn.

The main fact to point out is that this is enough information, in the 2×2 case, to tell us that the eigenvalues have to be real and distinct. Since their product is a negative number, we can eliminate the other two options. If we have two complex roots, they must be of the form $x + iy$ and $x - iy$, and so the product is

$$(x + iy)(x - iy) = x^2 + ixy - ixy - i^2y^2 = x^2 + y^2$$

which is always positive, no matter what x and y are. Similarly, if we have a repeated eigvalue, the product will be that number squared, which is also positive. Therefore, if the determinant of a 2×2 matrix is negative, the eigenvalues must be real and distinct, with one being positive and one negative (otherwise the product can not be negative). These facts will be important when we start to analyze the solutions to systems of differential equations in Chapter **??**.

Example 97. *What can be said about the eigenvalues of the matrix*

$$A = \begin{bmatrix} 0 & -1 & 0 \\ 2 & 2 & 0 \\ -7 & -3 & -1 \end{bmatrix}?$$

Solution: We can find the same information as the previous example. The trace of A is 1, and the determinant, by cofactor expansion along column 3, is $(-1)(0+2) = -2$. Therefore, the sum of the *three* eigenvalues is 1, and the product of them is -2. We don't actually have enough information here to determine what the eigenvalues are. The issue is that with three eigenvalues, there are many different ways to get to a product being negative. There could be three negative eigenvalues, two positive and one negative, or one negative real with two complex eigenvalues. However, the one thing we do know for sure is that there must be one negative real eigenvalue. For this particular example, we can compute that the eigenvalues are -1, $1 + i$, and $1 - i$, so we did end up in the complex case.

Exercise 492 *Imagine that we have a 3×3 matrix with a positive determinant (it doesn't matter what the trace is). Think about all the scenarios and verify that at least one eigenvalue must be real and positive for this to happen.*

Extension of Previous Theorem

With all of the new definitions and properties that have been stated, we can add a few more equivalent statements to Theorem 20.

Theorem 28. *Let A be an $n \times n$ matrix. The following are equivalent:*

(a) *A is invertible.*

(b) $\det(A) \neq 0$.

(g) *The rank of A is n.*

(h) *The rows of A are linearly independent.*

(i) *The nullity of the matrix is 0.*

(j) *None of the eigenvalues of A are 0, or equivalently, the product of the eigenvalues of A is non-zero.*

(k) *The columns of A are a basis of $\mathbb{R}^n$.*

(l) *The rows of A are a basis of $\mathbb{R}^n$.*

Proof Most of these follow from the components of Theorem 20. If A is invertible, then we know that the columns are linearly independent. But there are n columns, so that number must be the rank. This implies that the rows are linearly independent, and if the rank plus the nullity must be n, we must have the nullity equal to zero. On that same train of thought, if we have n linearly independent vectors in $\mathbb{R}^n$, then they must be a basis, giving (k) and (l). Finally, since the determinant is the product of the eigenvalues, if the determinant is non-zero, that implies fact (j). ■

`linearAlgebra/practice/advancedLA-practice1.tex`

Practice for Advanced Linear Algebra

Why?

Exercise 493 *For the following matrices, find a basis for the kernel (nullspace).*

a) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 5 \\ 1 & 1 & -4 \end{bmatrix}$ b) $\begin{bmatrix} 2 & -1 & -3 \\ 4 & 0 & -4 \\ -1 & 1 & 2 \end{bmatrix}$ c) $\begin{bmatrix} -4 & 4 & 4 \\ -1 & 1 & 1 \\ -5 & 5 & 5 \end{bmatrix}$ d) $\begin{bmatrix} -2 & 1 & 1 & 1 \\ -4 & 2 & 2 & 2 \\ 1 & 0 & 4 & 3 \end{bmatrix}$

Exercise 494 *For the following matrices, find a basis for the kernel (nullspace).*

a) $\begin{bmatrix} 2 & 6 & 1 & 9 \\ 1 & 3 & 2 & 9 \\ 3 & 9 & 0 & 9 \end{bmatrix}$ b) $\begin{bmatrix} 2 & -2 & -5 \\ -1 & 1 & 5 \\ -5 & 5 & -3 \end{bmatrix}$ c) $\begin{bmatrix} 1 & -5 & -4 \\ 2 & 3 & 5 \\ -3 & 5 & 2 \end{bmatrix}$ d) $\begin{bmatrix} 0 & 4 & 4 \\ 0 & 1 & 1 \\ 0 & 5 & 5 \end{bmatrix}$

Exercise 495 *Suppose a 5×5 matrix A has rank 3. What is the nullity?*

Exercise 496 *Consider a square matrix A, and suppose that $\vec{x}$ is a nonzero vector such that $A\vec{x} = \vec{0}$. What does the Fredholm alternative say about invertibility of A?*

Exercise 497 *Compute the rank of the matrix A below.*

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

What does this tell you about the invertibility of A? How about the solutions to $A\vec{x} = \vec{0}$?

Exercise 498 *Compute the rank of the matrix A below.*

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

What does this tell you about the invertibility of A? How about the solutions to $A\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$?

Exercise 499 *Consider*

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

If the nullity of this matrix is 2, fill in the question marks. Hint: What is the rank?

Author(s): Matthew Charnley and Jason Nowell

Exercise 500 *Suppose the column space of a 9×5 matrix A of dimension 3. Find*

a) *Rank of A.*

b) *Nullity of A.*

c) *Dimension of the row space of A.*

d) *Dimension of the nullspace of A.*

e) *Size of the maximum subset of linearly independent rows of A.*

Exercise 501 *Compute the rank of the given matrices*

a) $\begin{bmatrix} 6 & 3 & 5 \\ 1 & 4 & 1 \\ 7 & 7 & 6 \end{bmatrix}$ b) $\begin{bmatrix} 5 & -2 & -1 \\ 3 & 0 & 6 \\ 2 & 4 & 5 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 2 & 4 & 6 \end{bmatrix}$

Exercise 502 *Compute the rank of the given matrices*

a) $\begin{bmatrix} 7 & -1 & 6 \\ 7 & 7 & 7 \\ 7 & 6 & 2 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}$ c) $\begin{bmatrix} 0 & 3 & -1 \\ 6 & 3 & 1 \\ 4 & 7 & -1 \end{bmatrix}$

Exercise 503 *For the matrices in* Exercise *, find a linearly independent set of row vectors that span the row space (they don't need to be rows of the matrix).*

Exercise 504 *For the matrices in* Exercise *, find a linearly independent set of columns that span the column space. That is, find the pivot columns of the matrices.*

Exercise 505 *For the matrices in* Exercise *, find a linearly independent set of row vectors that span the row space (they don't need to be rows of the matrix).*

Exercise 506 *For the matrices in* Exercise *, find a linearly independent set of columns that span the column space. That is, find the pivot columns of the matrices.*

Exercise 507 *Compute the rank of the matrix*

$$\begin{bmatrix} 10 & -2 & 11 & -7 \\ -5 & -2 & -5 & 5 \\ 1 & 0 & -4 & -4 \\ 1 & 2 & 2 & -1 \end{bmatrix}$$

Exercise 508 *Compute the rank of the matrix*

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

Exercise 509 *Find a linearly independent subset of the following vectors that has the same span.*

$$\begin{bmatrix}-1\\1\\2\end{bmatrix},\quad\begin{bmatrix}2\\-2\\-4\end{bmatrix},\quad\begin{bmatrix}-2\\4\\1\end{bmatrix},\quad\begin{bmatrix}-1\\3\\-2\end{bmatrix}$$

Exercise 510 *Find a linearly independent subset of the following vectors that has the same span.*

$$\begin{bmatrix}0\\0\\0\end{bmatrix},\quad\begin{bmatrix}3\\1\\-5\end{bmatrix},\quad\begin{bmatrix}0\\3\\-1\end{bmatrix},\quad\begin{bmatrix}-3\\2\\4\end{bmatrix}$$

Exercise 511 *For the following sets of vectors, determine if the set is linearly independent. Then find a basis for the subspace spanned by the vectors, and find the dimension of the subspace.*

a) $\begin{bmatrix}1\\1\\1\end{bmatrix},\ \begin{bmatrix}-1\\-1\\-1\end{bmatrix}$

b) $\begin{bmatrix}1\\0\\5\end{bmatrix},\ \begin{bmatrix}0\\1\\0\end{bmatrix},\ \begin{bmatrix}0\\-1\\0\end{bmatrix}$

c) $\begin{bmatrix}-4\\-3\\5\end{bmatrix},\ \begin{bmatrix}2\\3\\3\end{bmatrix},\ \begin{bmatrix}2\\0\\2\end{bmatrix}$

d) $\begin{bmatrix}1\\3\\0\end{bmatrix},\ \begin{bmatrix}0\\2\\2\end{bmatrix},\ \begin{bmatrix}-1\\-1\\2\end{bmatrix}$

e) $\begin{bmatrix}1\\3\end{bmatrix},\ \begin{bmatrix}0\\2\end{bmatrix},\ \begin{bmatrix}-1\\-1\end{bmatrix}$

f) $\begin{bmatrix}3\\1\\3\end{bmatrix},\ \begin{bmatrix}2\\4\\-4\end{bmatrix},\ \begin{bmatrix}-5\\-5\\-2\end{bmatrix}$

Exercise 512 *For the following sets of vectors, determine if the set is linearly independent. Then find a basis for the subspace spanned by the vectors, and find the dimension of the subspace.*

a) $\begin{bmatrix}1\\2\end{bmatrix},\ \begin{bmatrix}1\\1\end{bmatrix}$

b) $\begin{bmatrix}1\\1\\1\end{bmatrix},\ \begin{bmatrix}2\\2\\2\end{bmatrix},\ \begin{bmatrix}1\\1\\2\end{bmatrix}$

c) $\begin{bmatrix}5\\3\\1\end{bmatrix},\ \begin{bmatrix}5\\-1\\5\end{bmatrix},\ \begin{bmatrix}-1\\3\\-4\end{bmatrix}$

d) $\begin{bmatrix}2\\2\\4\end{bmatrix},\ \begin{bmatrix}2\\2\\3\end{bmatrix},\ \begin{bmatrix}4\\4\\-3\end{bmatrix}$

e) $\begin{bmatrix}1\\0\end{bmatrix},\ \begin{bmatrix}2\\0\end{bmatrix},\ \begin{bmatrix}3\\0\end{bmatrix}$

f) $\begin{bmatrix}1\\0\\0\end{bmatrix},\ \begin{bmatrix}2\\0\\0\end{bmatrix},\ \begin{bmatrix}0\\1\\2\end{bmatrix}$

Exercise 513 *Suppose that X is the set of all the vectors of $\mathbb{R}^3$ whose third component is zero. Is X a subspace? And if so, find a basis and the dimension.*

Exercise 514 *Consider a set of 3 component vectors.*

a) *How can it be shown if these vectors are linearly independent?*

b) *Can a set of 4 of these 3 component vectors be linearly independent? Explain your answer.*

c) *Can a set of 2 of these 3 component vectors be linearly independent? Explain.*

d) *How would it be shown if these vectors make up a spanning set for all 3 component vectors?*

e) *Can 4 vectors be a spanning set? Explain.*

f) *Can 2 vectors be a spanning set? Explain.*

Exercise 515 *Consider the vectors*

$$\vec{v}_1 = \begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}.$$

Let A be the matrix with these vectors as columns and $\vec{b}$ the vector $[1\ 0\ 0]$.

a) *Compute the rank of A to determine how many of these vectors are linearly independent.*

b) *Determine if $\vec{b}$ is in the span of the given vectors by using row reduction to try to solve $A\vec{x} = \vec{b}$.*

c) *Look at the columns of the row-reduced form of A. Is $\vec{b}$ in the span of those vectors?*

d) *What do these last two parts tell you about the span of the columns of a matrix, and the span of the columns of the row-reduced matrix?*

e) *Now, build a matrix D with these vectors as rows. Row-reduce this matrix to get a matrix D_2.*

f) *Is $\vec{b}$ in the span of the rows of D_2? You can't check this in using the matrix form; instead, just brute force it based on the form of D_2. What does this potentially say about the span of the rows of D and the rows of D_2?*

Exercise 516 *Complete* Exercise *with*

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} -6 \\ 2 \\ 3 \\ -1 \end{bmatrix} \quad \begin{bmatrix} -13 \\ 3 \\ 1 \\ 1 \end{bmatrix} \quad \vec{v}_4 \begin{bmatrix} 11\ -1 \\ -5 \\ -1 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Exercise 517 *Compute the inverse of the given matrices*

a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}$

Exercise 518 *Compute the inverse of the given matrices*

a) $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ c) $\begin{bmatrix} 2 & 4 & 0 \\ 2 & 2 & 3 \\ 2 & 4 & 1 \end{bmatrix}$

Exercise 519 *By computing the inverse, solve the following systems for $\vec{x}$.*

a) $\begin{bmatrix} 4 & 1 \\ -1 & 3 \end{bmatrix} \vec{x} = \begin{bmatrix} 13 \\ 26 \end{bmatrix}$ b) $\begin{bmatrix} 3 & 3 \\ 3 & 4 \end{bmatrix} \vec{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

Exercise 520 *By computing the inverse, solve the following systems for $\vec{x}$.*

a) $\begin{bmatrix} -1 & 1 \\ 3 & 3 \end{bmatrix} \vec{x} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ b) $\begin{bmatrix} 2 & 7 \\ 1 & 6 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

Exercise 521 *For each of the following matrices below:*

a) *Compute the trace and determinant of the matrix, and*

b) *Find the eigenvalues of the matrix and verify that the trace is the sum of the eigenvalues and the determinant is the product.*

$$(i)\ \begin{bmatrix} -4 & 2 \\ -9 & 5 \end{bmatrix} \qquad (ii)\ \begin{bmatrix} 2 & -3 \\ 6 & -4 \end{bmatrix} \qquad (iii)\ \begin{bmatrix} -10 & -12 \\ 6 & 8 \end{bmatrix}. \qquad (iv)\ \begin{bmatrix} -7 & -9 \\ 1 & -1 \end{bmatrix}$$

Exercise 522 *For each of the following matrices below:*

a) *Compute the trace and determinant of the matrix, and*

b) *Find the eigenvalues of the matrix and verify that the trace is the sum of the eigenvalues and the determinant is the product.*

$$(i)\ \begin{bmatrix} -1 & -16 & -4 \\ 1 & 6 & 1 \\ -2 & -4 & 1 \end{bmatrix} \qquad (ii)\ \begin{bmatrix} 1 & 2 & 0 \\ -12 & -13 & -4 \\ 16 & 14 & 3 \end{bmatrix} \qquad (iii)\ \begin{bmatrix} 10 & -7 & -14 \\ 0 & 5 & 6 \\ 7 & -8 & -14 \end{bmatrix}$$

Nonhomogeneous systems

We discuss Nonhomogeneous systems

Now, we want to take a look at solving non-homogeneous linear systems. As discussed previously, the process here is the same as it was for second order non-homogeneous equations. We can solve the homogeneous equation and then need one particular solution to the non-homogeneous problem. Adding these together gives the general solution to the non-homogeneous problem, where we can pick constants to meet an initial condition if it is given. This section here will focus on a variety of methods to find this particular solution.

First order constant coefficient

Diagonalization

Diagonalization is a linear algebra-based process for adjusting a matrix into one that is diagonal. In order to see why this might be helpful in the process of solving non-homogeneous systems, or generating a particular solution to the non-homogeneous system, let's start by looking at a problem with a diagonal matrix to see how we could solve it.

Example 98. *Find the general solution of the non-homogeneous system*

$$\vec{x}'(t) = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \vec{x} + \begin{bmatrix} e^{2t} \\ e^{-t} \end{bmatrix}.$$

Solution: If we write this system out in components, we get

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} e^{2t} \\ e^{-t} \end{bmatrix},$$

or

$$x_1' = x_1 + e^{2t} \qquad x_2' = 3x_2 + e^{-t}.$$

These are two completely separated, or *decoupled* equations. We can solve each of these via first-order integrating factor methods. For the first, we get

$$\begin{aligned} x_1' - x_1 &= e^{2t} \\ (e^{-t}x_1)' &= e^{t} \\ e^{-t}x_1 &= e^{t} + C_1 \\ x_1(t) &= e^{2t} + C_1 e^{t} \end{aligned}$$

and for the second, we see that

$$\begin{aligned} x_2' - 3x_2 &= e^{-t} \\ (e^{-3t}x_2)' &= e^{2t} \\ e^{-3t}x_2 &= \frac{1}{2}e^{2t} + C_2 \\ x_2(t) &= \frac{1}{2}e^{-t} + C_2 e^{3t} \end{aligned}.$$

Therefore, the solution to this system is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} e^{2t} + C_1 e^{t} \\ \frac{1}{2}e^{-t} + C_2 e^{3t} \end{bmatrix}$$

Learning outcomes: Use the eigenvector decomposition or diagonalization to solve non-homogeneous systems Use undetermined coefficients to solve non-homogeneous systems Use variation of parameters and fundamental matrices of solutions to solve non-homogeneous systems.

Author(s): Matthew Charnley and Jason Nowell

or, rewriting in a different form,

$$\vec{x}(t) = \begin{bmatrix} e^{2t} \\ \frac{1}{2}e^{-t} \end{bmatrix} + C_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^t + C_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{3t}.$$

Therefore, if we have a non-homogeneous system with a diagonal matrix, then we can separate the decoupled equations, solve them individually, and put them back together into a full solution. In this particular case, the eigenvectors of A were $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and so the standard basis vectors were the directions in which A acts like a scalar. When the eigenvectors are not the standard basis vectors, we need to take them into account in order to use this method.

Take the equation

$$\vec{x}'(t) = A\vec{x}(t) + \vec{f}(t). \tag{33}$$

Assume A has n linearly independent eigenvectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Build the matrices

$$E = [\vec{v}_1 \mid \vec{v}_2 \mid \cdots \mid \vec{v}_n] \qquad D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

that is, E is the matrix with the eigenvectors as columns, and D is a diagonal matrix with the eigenvalues on the diagonal in the same order as the eigenvectors are put into E. Since we have n eigenvectors, both of these are $n \times n$ square matrices. It is a fact from linear algebra that

$$A = EDE^{-1} \quad \text{or} \quad D = E^{-1}AE.$$

Exercise 523 *For the matrix*

$$A = \begin{bmatrix} 6 & 2 \\ -4 & 0 \end{bmatrix}$$

compute the matrices E and D and verify that $EDE^{-1} = A$.

With this tool in hand, we look to approach our non-homogeneous system. We would like for the system to use the matrix D instead of the matrix A, because that is decoupled and we can solve it directly. To do this, we define a new unknown function $\vec{y}$ by the relation $\vec{x} = E\vec{y}$. If we plug this into (33), we get

$$E\vec{y}'(t) = AE\vec{y}(t) + \vec{f}(t).$$

Using the relation for A and the fact that E is a constant matrix, we get that

$$E\vec{y}'(t) = EDE^{-1}E\vec{y}(t) + \vec{f}(t) = ED\vec{y}(t) + \vec{f}(t).$$

If we multiply both sides of this equation by E^{-1}, we get

$$\vec{y}'(t) = D\vec{y}(t) + E^{-1}\vec{f}(t)$$

and this is now a decoupled system of equations. Once we compute $E^{-1}\vec{f}(t)$, we can then solve this directly because it is based on a decoupled system of differential equations to solve for the solution $\vec{y}$. Once we have $\vec{y}$, we can compute $\vec{x}$ as $\vec{x} = E\vec{y}$ to get our solution.

Example 99. *Let $A = \left[\begin{smallmatrix} 1 & 3 \\ 3 & 1 \end{smallmatrix}\right]$. Solve $\vec{x}' = A\vec{x} + \vec{f}$ where $\vec{f}(t) = \left[\begin{smallmatrix} 2e^t \\ 2t \end{smallmatrix}\right]$ for $\vec{x}(0) = \left[\begin{smallmatrix} 3/16 \\ -5/16 \end{smallmatrix}\right]$.*

Solution: The first step in this process is always to find the eigenvalues and eigenvectors of the coefficient matrix. We do this in the standard way

$$\det(A - \lambda I) = (1-\lambda)(1-\lambda) - (3)(3) = \lambda^2 - 2\lambda + 1 - 9 = \lambda^2 - 2\lambda - 8.$$

Since this factors as $(\lambda+2)(\lambda-4)$, the eigenvalues are -2 and 4. Using these (exercise!) we can show that the corresponding eigenvectors are $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$ for $\lambda = -2$ and $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ for $\lambda = 4$. Therefore, the general solution to the homogeneous problem is

$$\vec{x}(t) = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{4t}.$$

Now that we have this solution, we can work to solve the non-homogeneous problem. To do this, we form the matrices

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

and, using the fact that for a 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

we can compute E^{-1} as

$$E^{-1} = \frac{1}{(1)(1) - (1)(-1)} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}.$$

As an aside, we can check that $A = EDE^{-1}$ to make sure that we did this right.

$$\begin{aligned} EDE^{-1} &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} = A. \end{aligned}$$

Thus, we can proceed. From the general process of diagonalization, we know that the system we need to solve is

$$\vec{y}' = D\vec{y} + E^{-1}\vec{f} = \begin{bmatrix} -2 & 0 \\ 0 & 4 \end{bmatrix} \vec{y} + \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2e^t \\ 2t \end{bmatrix}$$

for $\vec{y} = E^{-1}\vec{x}$, or $\vec{y}$ defined by $\vec{x} = E\vec{y}$. Computing the non-homogenoeus term gives

$$\begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2e^t \\ 2t \end{bmatrix} = \begin{bmatrix} e^t - t \\ e^t + t \end{bmatrix}$$

so that we can now decouple the system

$$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} e^t - t \\ e^t + t \end{bmatrix}$$

into two separate first-order equations that we can solve

$$y_1' = -2y_1 + e^t - t \qquad y_2' = 4y_2 + e^t + t$$

by normal first-order integrating factor methods. For the y_1 equation, we want to use an integrating factor of e^{2t} to solve it as

$$\begin{aligned} y_1' + 2y_1 &= e^t - t \\ e^{2t}y_1' + 2e^{2t}y_1 &= e^{3t} - te^{2t} \\ (e^{2t}y_1)' &= e^{3t} - te^{2t} \\ e^{2t}y_1 &= \int e^{3t} - te^{2t}\, dt = \frac{1}{3}e^{3t} - \frac{1}{2}te^{2t} + \frac{1}{4}e^{2t} + C_1 \\ y_1 &= \frac{1}{3}e^t - \frac{1}{2}t + \frac{1}{4} + C_1e^{-2t}. \end{aligned}$$

For the second, we need the integrating factor e^{-4t} to solve

$$\begin{aligned} y_2' - 4y_2 &= e^t + t \\ e^{-4t}y_2' - 4e^{-4t}y_2 &= e^{-3t} + te^{-4t} \\ e^{-4t}y_2 &= \int e^{-3t} + te^{-4t}\, dt = -\frac{1}{3}e^{-3t} - \frac{1}{4}te^{-4t} - \frac{1}{16}e^{-4t} + C_2 \\ y_2 &= -\frac{1}{3}e^t - \frac{1}{4}t - \frac{1}{16} + C_2e^{4t}. \end{aligned}$$

Therefore, we have the vector solution

$$\vec{y}(t) = \begin{bmatrix} \frac{1}{3}e^t - \frac{1}{2}t + \frac{1}{4} + C_1 e^{-2t} \\ -\frac{1}{3}e^t - \frac{1}{4}t - \frac{1}{16} + C_2 e^{4t} \end{bmatrix}.$$

To get to the actual solution $\vec{x}$, we need to multiply this solution by the matrix E

$$\begin{aligned} \vec{x} = E\vec{y} = x \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} & \begin{bmatrix} \frac{1}{3}e^t - \frac{1}{2}t + \frac{1}{4} + C_1 e^{-2t} \\ -\frac{1}{3}e^t - \frac{1}{4}t - \frac{1}{16} + C_2 e^{4t} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{3}e^t - \frac{1}{2}t + \frac{1}{4} + C_1 e^{-2t} + (-\frac{1}{3}e^t - \frac{1}{4}t - \frac{1}{16} + C_2 e^{4t}) \\ -(\frac{1}{3}e^t - \frac{1}{2}t + \frac{1}{4} + C_1 e^{-2t}) + (-\frac{1}{3}e^t - \frac{1}{4}t - \frac{1}{16} + C_2 e^{4t}) \end{bmatrix} \\ &= \begin{bmatrix} -\frac{3}{4}t + \frac{3}{16} + C_1 e^{-2t} + C_2 e^{-4t} \\ -\frac{2}{3}e^t - \frac{1}{4}t - \frac{5}{16} - C_1 e^{-2t} + C_2 e^{4t} \end{bmatrix} \end{aligned}$$

which is a valid way to write the general solution. We can also write this solution in the form

$$\vec{x}(t) = \begin{bmatrix} 0 \\ -\frac{2}{3} \end{bmatrix} e^t + \begin{bmatrix} -\frac{3}{4} \\ -\frac{1}{4} \end{bmatrix} t + \begin{bmatrix} \frac{3}{16} \\ -\frac{5}{16} \end{bmatrix} + C_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-2t} + C_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{4t}$$

and we see that the general solution to the homogeneous problem shows up at the end of this solution.

Finally, we need to satisfy the initial conditions. If we plug in $t = 0$, we get

$$\vec{x}(0) = \begin{bmatrix} 0 \\ -\frac{2}{3} \end{bmatrix} + 0 + \begin{bmatrix} \frac{3}{16} \\ -\frac{5}{16} \end{bmatrix} + C_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + C_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3/16 \\ -5/16 \end{bmatrix}.$$

Rearranging this expression gives the two equations

$$C_1 + C_2 = 0 \qquad -C_1 + C_2 = \frac{2}{3}$$

which has solution $C_1 = -1/3$ and $C_2 = 1/3$. Therefore, the solution to the initial value problem is

$$\vec{x}(t) = \begin{bmatrix} 0 \\ -\frac{2}{3} \end{bmatrix} e^t + \begin{bmatrix} -\frac{3}{4} \\ -\frac{1}{4} \end{bmatrix} t + \begin{bmatrix} \frac{3}{16} \\ -\frac{5}{16} \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-2t} + \frac{1}{3} \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{4t}.$$

Another way to view this process is by thinking about it as eigenvector decomposition. (This approach is not necessary on a first reading. The next new information starts at the undetermined coefficients section.) The eigenvectors of A are the directions in which the matrix A basically acts like a scalar. If we can solve the differential equation in those directions, then it acts like a scalar equation, which we know how to solve. We can then reorient everything to get back to our original solution.

Again, we start with the equation

$$\vec{x}'(t) = A\vec{x}(t) + \vec{f}(t) \tag{34}$$

and assume A has n linearly independent eigenvectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ Write

$$\vec{x}(t) = \vec{v}_1\, \xi_1(t) + \vec{v}_2\, \xi_2(t) + \cdots + \vec{v}_n\, \xi_n(t). \tag{35}$$

That is, we wish to write our solution as a linear combination of eigenvectors of A. If we solve for the scalar functions ξ_1 through ξ_n, we have our solution $\vec{x}$. Let us decompose $\vec{f}$ in terms of the eigenvectors as well. We wish to write

$$\vec{f}(t) = \vec{v}_1\, g_1(t) + \vec{v}_2\, g_2(t) + \cdots + \vec{v}_n\, g_n(t). \tag{36}$$

That is, we wish to find g_1 through g_n that satisfy (36). Since all the eigenvectors are independent, the matrix $E = [\,\vec{v}_1 \quad \vec{v}_2 \quad \cdots \quad \vec{v}_n\,]$ is invertible. Write the equation (36) as $\vec{f} = E\vec{g}$, where the components of $\vec{g}$ are the functions g_1 through g_n. Then $\vec{g} = E^{-1}\vec{f}$. Hence it is always possible to find $\vec{g}$ when there are n linearly independent eigenvectors.

We plug (35) into (34), and note that $A\vec{v}_k = \lambda_k \vec{v}_k$:

$$\begin{aligned}
\overbrace{\vec{v}_1\xi_1' + \vec{v}_2\xi_2' + \cdots + \vec{v}_n\xi_n'}^{\vec{x}'} &= \overbrace{A\left(\vec{v}_1\xi_1 + \vec{v}_2\xi_2 + \cdots + \vec{v}_n\xi_n\right)}^{A\vec{x}} + \overbrace{\vec{v}_1 g_1 + \vec{v}_2 g_2 + \cdots + \vec{v}_n g_n}^{\vec{f}} \\
&= A\vec{v}_1\xi_1 + A\vec{v}_2\xi_2 + \cdots + A\vec{v}_n\xi_n + \vec{v}_1 g_1 + \vec{v}_2 g_2 + \cdots + \vec{v}_n g_n \\
&= \vec{v}_1\lambda_1\xi_1 + \vec{v}_2\lambda_2\xi_2 + \cdots + \vec{v}_n\lambda_n\xi_n + \vec{v}_1 g_1 + \vec{v}_2 g_2 + \cdots + \vec{v}_n g_n \\
&= \vec{v}_1(\lambda_1\xi_1 + g_1) + \vec{v}_2(\lambda_2\xi_2 + g_2) + \cdots + \vec{v}_n(\lambda_n\xi_n + g_n).
\end{aligned}$$

If we identify the coefficients of the vectors $\vec{v}_1$ through $\vec{v}_n$, we get the equations

$$\begin{aligned}
\xi_1' &= \lambda_1\xi_1 + g_1, \\
\xi_2' &= \lambda_2\xi_2 + g_2, \\
&\ \ \vdots \\
\xi_n' &= \lambda_n\xi_n + g_n.
\end{aligned}$$

Each one of these equations is independent of the others. They are all linear first order equations and can easily be solved by the standard integrating factor method for single equations. That is, for the k^{th} equation we write

$$\xi_k'(t) - \lambda_k\xi_k(t) = g_k(t).$$

We use the integrating factor $e^{-\lambda_k t}$ to find that

$$\frac{d}{dt}\Big[\xi_k(t)\,e^{-\lambda_k t}\Big] = e^{-\lambda_k t} g_k(t).$$

We integrate and solve for ξ_k to get

$$\xi_k(t) = e^{\lambda_k t}\int e^{-\lambda_k t} g_k(t)\,dt + C_k e^{\lambda_k t}.$$

If we are looking for just any particular solution, we can set C_k to be zero. If we leave these constants in, we get the general solution. Write $\vec{x}(t) = \vec{v}_1\xi_1(t) + \vec{v}_2\xi_2(t) + \cdots + \vec{v}_n\xi_n(t)$, and we are done.

As always, it is perhaps better to write these integrals as definite integrals. Suppose that we have an initial condition $\vec{x}(0) = \vec{b}$. Take $\vec{a} = E^{-1}\vec{b}$ to find $\vec{b} = \vec{v}_1 a_1 + \vec{v}_2 a_2 + \cdots + \vec{v}_n a_n$, just like before. Then if we write

$$\xi_k(t) = e^{\lambda_k t}\int_0^t e^{-\lambda_k s} g_k(s)\,ds + a_k e^{\lambda_k t},$$

we get the particular solution $\vec{x}(t) = \vec{v}_1\xi_1(t) + \vec{v}_2\xi_2(t) + \cdots + \vec{v}_n\xi_n(t)$ satisfying $\vec{x}(0) = \vec{b}$, because $\xi_k(0) = a_k$.

Let us remark that the technique we just outlined is the eigenvalue method applied to nonhomogeneous systems. If a system is homogeneous, that is, if $\vec{f} = \vec{0}$, then the equations we get are $\xi_k' = \lambda_k\xi_k$, and so $\xi_k = C_k e^{\lambda_k t}$ are the solutions and that's precisely what we got in § **??**.

Example 100. *(Same as the previous example) Let* $A = \left[\begin{smallmatrix} 1 & 3 \\ 3 & 1 \end{smallmatrix}\right]$. *Solve* $\vec{x}' = A\vec{x} + \vec{f}$ *where* $\vec{f}(t) = \left[\begin{smallmatrix} 2e^t \\ 2t \end{smallmatrix}\right]$ *for* $\vec{x}(0) = \begin{bmatrix} 3/16 \\ -5/16 \end{bmatrix}$.

Solution: The eigenvalues of A are -2 and 4 and corresponding eigenvectors are $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ respectively. We write down the matrix E of the eigenvectors and compute its inverse (using the inverse formula for 2×2 matrices)

$$E = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \qquad E^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

We are looking for a solution of the form $\vec{x} = \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]\xi_1 + \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]\xi_2$. We first need to write $\vec{f}$ in terms of the eigenvectors. That is we wish to write $\vec{f} = \left[\begin{smallmatrix} 2e^t \\ 2t \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right] g_1 + \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right] g_2$. Thus

$$\begin{bmatrix} g_1 \\ g_2 \end{bmatrix} = E^{-1}\begin{bmatrix} 2e^t \\ 2t \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2e^t \\ 2t \end{bmatrix} = \begin{bmatrix} e^t - t \\ e^t + t \end{bmatrix}.$$

So $g_1 = e^t - t$ and $g_2 = e^t + t$.

We further need to write $\vec{x}(0)$ in terms of the eigenvectors. That is, we wish to write $\vec{x}(0) = \left[\begin{smallmatrix} 3/16 \\ -5/16 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right] a_1 + \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right] a_2$. Hence

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = E^{-1} \begin{bmatrix} 3/16 \\ -5/16 \end{bmatrix} = \begin{bmatrix} 1/4 \\ -1/16 \end{bmatrix}.$$

So $a_1 = 1/4$ and $a_2 = -1/16$. We plug our $\vec{x}$ into the equation and get

$$\begin{aligned}
\overbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix} \xi_1' + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \xi_2'}^{\vec{x}'} &= \overbrace{A \begin{bmatrix} 1 \\ -1 \end{bmatrix} \xi_1 + A \begin{bmatrix} 1 \\ 1 \end{bmatrix} \xi_2}^{A\vec{x}} + \overbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix} g_1 + \begin{bmatrix} 1 \\ 1 \end{bmatrix} g_2}^{\vec{f}} \\
&= \begin{bmatrix} 1 \\ -1 \end{bmatrix} (-2\xi_1) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} 4\xi_2 + \begin{bmatrix} 1 \\ -1 \end{bmatrix} (e^t - t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} (e^t + t).
\end{aligned}$$

We get the two equations

$$\begin{aligned}
\xi_1' &= -2\xi_1 + e^t - t, & \text{where } \xi_1(0) = a_1 = \frac{1}{4}, \\
\xi_2' &= 4\xi_2 + e^t + t, & \text{where } \xi_2(0) = a_2 = \frac{-1}{16}.
\end{aligned}$$

We solve with integrating factor. Computation of the integral is left as an exercise to the student. You will need integration by parts.

$$\xi_1 = e^{-2t} \int e^{2t} \, (e^t - t) \, dt + C_1 e^{-2t} = \frac{e^t}{3} - \frac{t}{2} + \frac{1}{4} + C_1 e^{-2t}.$$

C_1 is the constant of integration. As $\xi_1(0) = 1/4$, then $1/4 = 1/3 + 1/4 + C_1$ and hence $C_1 = -1/3$. Similarly

$$\xi_2 = e^{4t} \int e^{-4t} \, (e^t + t) \, dt + C_2 e^{4t} = -\frac{e^t}{3} - \frac{t}{4} - \frac{1}{16} + C_2 e^{4t}.$$

As $\xi_2(0) = -1/16$ we have $-1/16 = -1/3 - 1/16 + C_2$ and hence $C_2 = 1/3$. The solution is

$$\vec{x}(t) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \underbrace{\left(\frac{e^t - e^{-2t}}{3} + \frac{1 - 2t}{4} \right)}_{\xi_1} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \underbrace{\left(\frac{e^{4t} - e^t}{3} - \frac{4t + 1}{16} \right)}_{\xi_2} = \begin{bmatrix} \frac{e^{4t} - e^{-2t}}{3} + \frac{3 - 12t}{16} \\ \frac{e^{-2t} + e^{4t} - 2e^t}{3} + \frac{4t - 5}{16} \end{bmatrix}.$$

That is, $x_1 = \dfrac{e^{4t} - e^{-2t}}{3} + \dfrac{3 - 12t}{16}$ and $x_2 = \dfrac{e^{-2t} + e^{4t} - 2e^t}{3} + \dfrac{4t - 5}{16}$.

Exercise 524 *Check that x_1 and x_2 solve the problem. Check both that they satisfy the differential equation and that they satisfy the initial conditions.*

Undetermined coefficients

The method of undetermined coefficients also works for systems. The only difference is that we use unknown vectors rather than just numbers. Same caveats apply to undetermined coefficients for systems as for single equations. This method does not always work for the same reasons that the corresponding method did not work for second order equations. We need to have a right-hand side of a proper form so that we can "guess" a solution of the correct form for the non-homogeneous solution. Furthermore, if the right-hand side is complicated, we have to solve for lots of variables. Each element of an unknown vector is an unknown number. In system of 3 equations with say say 4 unknown vectors (this would not be uncommon), we already have 12 unknown numbers to solve for. The method can turn into a lot of tedious work if done by hand. As the method is essentially the same as for single equations, let us just do an example.

Example 101. *Let $A = \left[\begin{smallmatrix} -1 & 0 \\ -2 & 1 \end{smallmatrix}\right]$. Find a particular solution of $\vec{x}' = A\vec{x} + \vec{f}$ where $\vec{f}(t) = \left[\begin{smallmatrix} e^t \\ t \end{smallmatrix}\right]$.*

Solution: Note that we can solve this system in an easier way (can you see how?), but for the purposes of the example, let us use the eigenvalue method plus undetermined coefficients. The eigenvalues of A are -1 and 1 and corresponding eigenvectors are $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$ respectively. Hence our complementary solution is

$$\vec{x}_c = c_1 \begin{bmatrix}1\\1\end{bmatrix} e^{-t} + c_2 \begin{bmatrix}0\\1\end{bmatrix} e^{t},$$

for some arbitrary constants c_1 and c_2.

We would want to guess a particular solution of

$$\vec{x} = \vec{a}e^t + \vec{b}t + \vec{d}.$$

However, something of the form $\vec{a}e^t$ appears in the complementary solution. Because we do not yet know if the vector $\vec{a}$ is a multiple of $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$, we do not know if a conflict arises. It is possible that there is no conflict, but to be safe we should also try $\vec{k}te^t$. Here we find the crux of the difference between a single equation and systems. We try *both* terms $\vec{a}e^t$ and $\vec{k}te^t$ in the solution, not just the term $\vec{k}te^t$. Therefore, we try

$$\vec{x} = \vec{a}e^t + \vec{k}te^t + \vec{b}t + \vec{d}.$$

Thus we have 8 unknowns. We write $\vec{a} = \left[\begin{smallmatrix}a_1\\a_2\end{smallmatrix}\right]$, $\vec{b} = \left[\begin{smallmatrix}b_1\\b_2\end{smallmatrix}\right]$, $\vec{k} = \left[\begin{smallmatrix}k_1\\k_2\end{smallmatrix}\right]$, and $\vec{d} = \left[\begin{smallmatrix}d_1\\d_2\end{smallmatrix}\right]$. We plug $\vec{x}$ into the equation. First let us compute $\vec{x}'$.

$$\vec{x}' = \left(\vec{a}+\vec{k}\right) e^t + \vec{k}te^t + \vec{b} = \begin{bmatrix}a_1+k_1\\a_2+k_2\end{bmatrix} e^t + \begin{bmatrix}k_1\\k_2\end{bmatrix} te^t + \begin{bmatrix}b_1\\b_2\end{bmatrix}.$$

Now $\vec{x}'$ must equal $A\vec{x}+\vec{f}$, which is

$$\begin{aligned}
A\vec{x}+\vec{f} &= A\vec{a}e^t + A\vec{k}te^t + A\vec{b}t + A\vec{d} + \vec{f} \\
&= \begin{bmatrix}-a_1\\-2a_1+a_2\end{bmatrix} e^t + \begin{bmatrix}-k_1\\-2k_1+k_2\end{bmatrix} te^t + \begin{bmatrix}-b_1\\-2b_1+b_2\end{bmatrix} t + \begin{bmatrix}-d_1\\-2d_1+d_2\end{bmatrix} + \begin{bmatrix}1\\0\end{bmatrix} e^t + \begin{bmatrix}0\\1\end{bmatrix} t \\
&= \begin{bmatrix}-a_1+1\\-2a_1+a_2\end{bmatrix} e^t + \begin{bmatrix}-k_1\\-2k_1+k_2\end{bmatrix} te^t + \begin{bmatrix}-b_1\\-2b_1+b_2+1\end{bmatrix} t + \begin{bmatrix}-d_1\\-2d_1+d_2\end{bmatrix}.
\end{aligned}$$

We identify the coefficients of e^t, te^t, t and any constant vectors in $\vec{x}'$ and in $A\vec{x}+\vec{f}$ to find the equations:

$$\begin{aligned}
a_1+k_1 &= -a_1+1, & 0 &= -b_1, \\
a_2+k_2 &= -2a_1+a_2, & 0 &= -2b_1+b_2+1, \\
k_1 &= -k_1, & b_1 &= -d_1, \\
k_2 &= -2k_1+k_2, & b_2 &= -2d_1+d_2.
\end{aligned}$$

We could write the 8×9 augmented matrix and start row reduction, but it is easier to just solve the equations in an ad hoc manner. Immediately we see that $k_1 = 0$, $b_1 = 0$, $d_1 = 0$. Plugging these back in, we get that $b_2 = -1$ and $d_2 = -1$. The remaining equations that tell us something are

$$\begin{aligned}
a_1 &= -a_1+1, \\
a_2+k_2 &= -2a_1+a_2.
\end{aligned}$$

So $a_1 = 1/2$ and $k_2 = -1$. Finally, a_2 can be arbitrary and still satisfy the equations. We are looking for just a single solution so presumably the simplest one is when $a_2 = 0$. Therefore,

$$\vec{x} = \vec{a}e^t + \vec{k}te^t + \vec{b}t + \vec{d} = \begin{bmatrix}1/2\\0\end{bmatrix} e^t + \begin{bmatrix}0\\-1\end{bmatrix} te^t + \begin{bmatrix}0\\-1\end{bmatrix} t + \begin{bmatrix}0\\-1\end{bmatrix} = \begin{bmatrix}\frac{1}{2}e^t\\-te^t-t-1\end{bmatrix}.$$

That is, $x_1 = \frac{1}{2}e^t$, $x_2 = -te^t - t - 1$. We would add this to the complementary solution to get the general solution of the problem. Notice that both $\vec{a}e^t$ and $\vec{k}te^t$ were really needed.

Exercise 525 *Check that x_1 and x_2 solve the problem. Try setting $a_2 = 1$ and check we get a solution as well. What is the difference between the two solutions we obtained (one with $a_2 = 0$ and one with $a_2 = 1$)?*

As you can see, other than the handling of conflicts, undetermined coefficients works exactly the same as it did for single equations. However, the computations can get out of hand pretty quickly for systems. The equation we considered was pretty simple.

First order variable coefficient

Variation of parameters

Just as for a single equation, there is the method of variation of parameters. This method works for any linear system, even if it is not constant coefficient, provided we somehow solve the associated homogeneous problem.

Suppose we have the equation

$$\vec{x}' = A(t)\,\vec{x} + \vec{f}(t). \tag{37}$$

Further, suppose we solved the associated homogeneous equation $\vec{x}' = A(t)\,\vec{x}$ and found a fundamental matrix solution $X(t)$. If we find separate, linearly independent solutions, this matrix $X(t)$ can be generated by putting these solutions as the columns of a matrix. The general solution to the associated homogeneous equation is $X(t)\vec{c}$ for a constant vector $\vec{c}$. Just like for variation of parameters for single equation we try the solution to the nonhomogeneous equation of the form

$$\vec{x}_p = X(t)\,\vec{u}(t),$$

where $\vec{u}(t)$ is a vector-valued function instead of a constant. We substitute $\vec{x}_p$ into (37) to obtain

$$\underbrace{X'(t)\,\vec{u}(t) + X(t)\,\vec{u}'(t)}_{\vec{x}_p'(t)} = \underbrace{A(t)\,X(t)\,\vec{u}(t)}_{A(t)\vec{x}_p(t)} + \vec{f}(t).$$

But $X(t)$ is a fundamental matrix solution to the homogeneous problem. So $X'(t) = A(t)X(t)$, and

$$\cancel{X'(t)\,\vec{u}(t)} + X(t)\,\vec{u}'(t) = \cancel{X'(t)\,\vec{u}(t)} + \vec{f}(t).$$

Hence $X(t)\,\vec{u}'(t) = \vec{f}(t)$. If we compute $[X(t)]^{-1}$, then $\vec{u}'(t) = [X(t)]^{-1}\,\vec{f}(t)$. We integrate to obtain $\vec{u}$ and we have the particular solution $\vec{x}_p = X(t)\,\vec{u}(t)$. Let us write this as a formula

$$\vec{x}_p = X(t) \int [X(t)]^{-1}\,\vec{f}(t)\,dt.$$

Example 102. *Find a particular solution to*

$$\vec{x}' = \frac{1}{t^2+1}\begin{bmatrix} t & -1 \\ 1 & t \end{bmatrix}\vec{x} + \begin{bmatrix} t \\ 1 \end{bmatrix}(t^2+1), \tag{38}$$

given that the general solution to the homogeneous problem

$$\vec{x}' = \frac{1}{t^2+1}\begin{bmatrix} t & -1 \\ 1 & t \end{bmatrix}\vec{x}$$

is

$$\vec{x}_c(t) = c_1\begin{bmatrix} 1 \\ t \end{bmatrix} + c_2\begin{bmatrix} -t \\ 1 \end{bmatrix}.$$

Solution: Here $A = \frac{1}{t^2+1}\left[\begin{smallmatrix} t & -1 \\ 1 & t \end{smallmatrix}\right]$ is most definitely not constant, so it's a good thing that we have the general solution to this system. From this, we can build the matrix $X(t)$ as

$$X = \begin{bmatrix} 1 & -t \\ t & 1 \end{bmatrix}$$

which is a fundamental matrix for this system and solves $X'(t) = A(t)X(t)$. Once we know the complementary solution we can find a solution to (38). First we find

$$[X(t)]^{-1} = \frac{1}{t^2+1}\begin{bmatrix} 1 & t \\ -t & 1 \end{bmatrix}.$$

Next we know a particular solution to (38) is

$$\begin{aligned}
\vec{x}_p &= X(t)\int [X(t)]^{-1}\vec{f}(t)\,dt \\
&= \begin{bmatrix} 1 & -t \\ t & 1 \end{bmatrix} \int \frac{1}{t^2+1}\begin{bmatrix} 1 & t \\ -t & 1 \end{bmatrix}\begin{bmatrix} t \\ 1 \end{bmatrix}(t^2+1)\,dt \\
&= \begin{bmatrix} 1 & -t \\ t & 1 \end{bmatrix} \int \begin{bmatrix} 2t \\ -t^2+1 \end{bmatrix} dt \\
&= \begin{bmatrix} 1 & -t \\ t & 1 \end{bmatrix} \begin{bmatrix} t^2 \\ -\frac{1}{3}t^3+t \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{3}t^4 \\ \frac{2}{3}t^3+t \end{bmatrix}.
\end{aligned}$$

Adding the complementary solution we find the general solution to (38):

$$\vec{x} = \begin{bmatrix} 1 & -t \\ t & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \frac{1}{3}t^4 \\ \frac{2}{3}t^3+t \end{bmatrix} = \begin{bmatrix} c_1 - c_2 t + \frac{1}{3}t^4 \\ c_2 + (c_1+1)\,t + \frac{2}{3}t^3 \end{bmatrix}.$$

Exercise 526 *Check that* $x_1 = \frac{1}{3}t^4$ *and* $x_2 = \frac{2}{3}t^3 + t$ *really solve* (38).

In the variation of parameters, we can obtain the general solution by adding in constants of integration. That is, we will add $X(t)\vec{c}$ for a vector of arbitrary constants. But that is precisely the complementary solution.

To conclude this section, we will solve one example using all three methods to be able to compare and contrast them. All of them have their benefits and drawbacks, and it's good to be able to do all three to be able to choose which to apply in a given circumstance.

Example 103. *Find the general solution to the system of differential equations*

$$\vec{x}' = \begin{bmatrix} -5 & -2 \\ 4 & 1 \end{bmatrix}\vec{x} + \begin{bmatrix} e^{2t}+1 \\ e^{2t}+3 \end{bmatrix}.$$

Solution: No matter which of the three methods we want to use to solve this problem, we always need the eigenvalues and eigenvectors of the coefficient matrix in order to find the general solution to the homogeneous problem. These are found by

$$\det(A-\lambda I) = (-5-\lambda)(1-\lambda) - (-2)(4) = \lambda^2 + 4\lambda - 5 + 8 = \lambda^2 + 4\lambda + 3.$$

This polynomial factors as $(\lambda+1)(\lambda+3)$ so the eigenvalues are -1 and -3. For $\lambda = -1$, the system we need to solve is

$$(A+I)\vec{v} = \begin{bmatrix} -4 & -2 \\ 4 & 2 \end{bmatrix}\vec{v} = \vec{0}$$

which can be solved by the vector $\vec{v} = \left[\begin{smallmatrix} 1 \\ -2 \end{smallmatrix}\right]$. For $\lambda = -3$, the system is

$$(A+3I)\vec{v} = \begin{bmatrix} -2 & -2 \\ 4 & 4 \end{bmatrix}\vec{v} = \vec{0}$$

which can be solved by the vector $\vec{v} = \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$. Therefore, the general solution to the homogeneous problem is

$$\vec{x}_c(t) = C_1\begin{bmatrix} 1 \\ -2 \end{bmatrix}e^{-t} + C_2\begin{bmatrix} 1 \\ -1 \end{bmatrix}e^{-3t}. \tag{39}$$

Now, we can divide into the different methods that we want to use to solve the non-homogeneous problem.

(a) Diagonalization. For this method, we need the matrices E and D defined by

$$E = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \qquad D = \begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix}$$

and can then compute E^{-1} as

$$E^{-1} = \frac{1}{(1)(-1) - (1)(-2)} \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix}.$$

We then compute

$$E^{-1}\vec{f} = \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} e^{2t} + 1 \\ e^{2t} + 3 \end{bmatrix} = \begin{bmatrix} -2e^{2t} - 4 \\ 3e^{2t} + 5 \end{bmatrix}$$

which gives rise to the decoupled system

$$\vec{y}' = \begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix} \vec{y} + \begin{bmatrix} -2e^{2t} - 4 \\ 3e^{2t} + 5 \end{bmatrix}$$

where $\vec{y}$ is defined by $\vec{x} = E\vec{y}$. We can solve for y_1 and y_2 using normal first-order methods:

$$\begin{aligned} y_1' + y_1 &= -2e^{2t} - 4 \\ (e^t y_1)' &= -2e^{3t} - 4e \\ e^t y_1 &= -\frac{2}{3}e^{3t} - 4e \\ y_1 &= -\frac{2}{3}e^{2t} - 4 \end{aligned}$$

$$\begin{aligned} y_2' + 3y_2 &= 3e^{2t} + 5 \\ (e^{3t} y_2)' &= 3e^{5t} + 5e^{3t} \\ e^{3t} y_2 &= \frac{3}{5}e^{5t} + \frac{5}{3}e^{3t} + C_2 \\ y_2 &= \frac{3}{5}e^{2t} + \frac{5}{3} + C_2 e^{-3t} \end{aligned}$$

Therefore, our solution for $\vec{y}$ is

$$\vec{y}(t) = \begin{bmatrix} -\frac{2}{3}e^{2t} - 4 + C_1 e^{-t} \\ \frac{3}{5}e^{2t} + \frac{5}{3} + C_2 e^{-3t} \end{bmatrix}$$

and by converting back to $\vec{x}$, we get

$$\begin{aligned} \vec{x}(t) = E\vec{y} &= \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} -\frac{2}{3}e^{2t} - 4 + C_1 e^{-t} \\ \frac{3}{5}e^{2t} + \frac{5}{3} + C_2 e^{-3t} \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{15}e^{2t} - \frac{7}{3} + C_1 e^{-t} + C_2 e^{-3t} \\ \frac{11}{15}e^{2t} + \frac{19}{3} - 2C_1 e^{-t} - C_2 e^{-3t} \end{bmatrix}. \end{aligned}$$

Or, rewriting in a different way,

$$\vec{x}(t) = \begin{bmatrix} -\frac{1}{15} \\ \frac{11}{15} \end{bmatrix} e^{2t} + \begin{bmatrix} -\frac{7}{3} \\ \frac{19}{3} \end{bmatrix} + C_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{-t} + C_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-3t}. \tag{40}$$

Notice how the general solution to the homogeneous equation (39) shows up at the end of this expression.

(b) Undetermined coefficients. Since the non-homogeneous part of our equation has terms of the form e^{2t} and constants, we should make a guess of the form

$$\vec{x}_p(t) = \vec{B}e^{2t} + \vec{D}.$$

We can plug this into our equation to get that

$$\vec{x}_p' = 2\vec{B}e^{2t} \tag{41}$$

and the right hand side of the equation is

$$\begin{bmatrix} -5 & -2 \\ 4 & 1 \end{bmatrix} \left(\vec{B} e^{2t} + \vec{D} \right) + \begin{bmatrix} e^{2t} + 1 \\ e^{2t} + 3 \end{bmatrix}.$$

Writing out $\vec{B}$ and $\vec{D}$ in components will give the right-hand side as

$$\begin{aligned} &\begin{bmatrix} -5b_1 - 2b_2 \\ 4b_1 + b_2 \end{bmatrix} e^{2t} + \begin{bmatrix} -5d_1 - 2d_2 \\ 4d_1 + d_2 \end{bmatrix} + \begin{bmatrix} e^{2t} + 1 \\ e^{2t} + 3 \end{bmatrix} \\ &= \begin{bmatrix} -5b_1 - 2b_2 + 1 \\ 4b_1 + b_2 + 1 \end{bmatrix} e^{2t} + \begin{bmatrix} -5d_1 - 2d_2 + 1 \\ 4d_1 + d_2 + 3 \end{bmatrix}. \end{aligned}$$

We can now set this equal to the left-hand side in (41) to get the vector equation

$$\begin{bmatrix} 2b_1 \\ 2b_2 \end{bmatrix} e^{2t} = \begin{bmatrix} -5b_1 - 2b_2 + 1 \\ 4b_1 + b_2 + 1 \end{bmatrix} e^{2t} + \begin{bmatrix} -5d_1 - 2d_2 + 1 \\ 4d_1 + d_2 + 3 \end{bmatrix}$$

and we can match up the terms on the left and right sides to get a system that we need to solve:

$$\begin{aligned} 2b_1 &= -5b_1 - 2b_2 + 1 \\ 2b_2 &= 4b_1 + b_2 + 1 \\ 0 &= -5d_1 - 2d_2 + 1 \\ 0 &= 4d_1 + d_2 + 3. \end{aligned}$$

Let's start with the b equations. Rearranging these gives

$$7b_1 + 2b_2 = 1 \qquad -4b_1 + b_2 = 1$$

Subtracting two copies of the second equation from the first gives $15b_1 = -1$ or $b_1 = -1/15$, which gives $b_2 = 1 + \dfrac{4}{15} = \dfrac{19}{13}$. Next, we can solve the d equations, which we can rearrange to give

$$5d_1 + 2d_2 = 1 \qquad 4d_1 + d_2 = -3$$

Subtracting two copies of the second equation from the first gives $-3d_1 = 7$ so $d_1 = -7/3$, leading to $d_2 = -3 - 4(-7/3) = 19/3$. Therefore, a solution to the non-homogeneous problem is

$$\vec{x}_p(t) = \begin{bmatrix} -\dfrac{1}{15} \\ \dfrac{19}{15} \end{bmatrix} e^{2t} + \begin{bmatrix} -\dfrac{7}{3} \\ \dfrac{19}{3} \end{bmatrix}$$

and so we can add in the homogeneous solution from (39) to get the full general solution as

$$\begin{bmatrix} -\dfrac{1}{13} \\ \dfrac{17}{13} \end{bmatrix} e^{2t} + \begin{bmatrix} -\dfrac{7}{3} \\ \dfrac{19}{3} \end{bmatrix} + C_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{-t} + C_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-3t}. \tag{42}$$

(c) Variation of Parameters. For this method, we write down the fundamental matrix $X(t)$ by combining the two basis solutions into a matrix, as

$$X(t) = \begin{bmatrix} e^{-t} & e^{-3t} \\ -2e^{-t} & -e^{-3t} \end{bmatrix}$$

and compute the inverse matrix as

$$X^{-1}(t) = \frac{1}{(e^{-t})(-e^{-3t}) - (e^{-3t})(-2e^{-t})} \begin{bmatrix} -e^{-3t} & -e^{-3t} \\ 2e^{-t} & e^{-t} \end{bmatrix} = \begin{bmatrix} -e^{t} & -e^{t} \\ 2e^{3t} & e^{3t} \end{bmatrix}.$$

We can then work out the components of the method of variation of parameters.

$$\begin{aligned} X(t)^{-1}\vec{f} &= \begin{bmatrix} -e^{t} & -e^{t} \\ 2e^{3t} & e^{3t} \end{bmatrix} \begin{bmatrix} e^{2t} + 1 \\ e^{2t} + 3 \end{bmatrix} \\ &= \begin{bmatrix} -e^{3t} - e^{t} - e^{3t} - 3e^{t} \\ 2e^{5t} + 2e^{3t} + e^{5t} + 3e^{3t} \end{bmatrix} \\ &= \begin{bmatrix} -2e^{3t} - 4e^{t} \\ 3e^{5t} + 5e^{3t} \end{bmatrix}. \end{aligned}$$

Integrating this expression gives

$$\int X(t)^{-1}\vec{f}\,dt = \begin{bmatrix} -\frac{2}{3}e^{3t} - 4e^{t} + C_1 \\ \frac{3}{5}e^{5t} + \frac{5}{3}e^{3t} + C_2 \end{bmatrix},$$

and so the general solution to this system is

$$\begin{aligned} X(t)\int X(t)^{-1}\vec{f}\,dt &= \begin{bmatrix} e^{-t} & e^{-3t} \\ -2e^{-t} & -e^{-3t} \end{bmatrix} \begin{bmatrix} -\frac{2}{3}e^{3t} - 4e^{t} + C_1 \\ \frac{3}{5}e^{5t} + \frac{5}{3}e^{3t} + C_2 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{2}{3}e^{2t} - 4 + C_1e^{-t} + \frac{3}{5}e^{2t} + \frac{5}{3} + C_2e^{-3t} \\ \frac{4}{3}e^{2t} + 8 - 2C_1e^{-t} - \frac{3}{5}e^{2t} - \frac{5}{3} - C_2e^{-3t} \end{bmatrix} \qquad (43) \\ &= \begin{bmatrix} -\frac{1}{15}e^{2t} - \frac{7}{3} + C_1e^{-t} + C_2e^{-3t} \\ \frac{11}{15}e^{2t} + \frac{19}{3} - 2C_1e^{-t} - C_2e^{-3t} \end{bmatrix}. \end{aligned}$$

Notice again that the homogeneous solution (39) shows up at the end of these terms, so we do not need to add it in at the end.

Comparing the solutions (40), (42), and (43), we see that the three solutions generated by these three methods are all the same.

For this previous example, we only found the general solution. If the solution to an initial value problem was needed, we would need to wait until the very end, once we have figured out the solution to the non-homogeneous problem and added in the solution to the homogeneous problem to determine the value of the constants to meet the initial condition.

`linearSystems/practice/nonHomogenSystems-practice1.tex`

Practice for Non-Homogeneous Systems

Why?

Exercise 527 *Find a particular solution to* $x' = x + 2y + 2t$, $y' = 3x + 2y - 4$,

a) *using diagonalization,*

b) *using undetermined coefficients.*

Exercise 528 *Find a particular solution to* $x' = 5x + 4y + t$, $y' = x + 8y - t$,

a) *using diagonalization,*

b) *using undetermined coefficients.*

Exercise 529 *Find the general solution to* $x' = 4x + y - 1$, $y' = x + 4y - e^t$,

a) *using diagonalization,*

b) *using undetermined coefficients.*

Exercise 530 *Find a particular solution to* $x' = y + e^t$, $y' = x + e^t$,

a) *using diagonalization,*

b) *using undetermined coefficients.*

Exercise 531 *Let* $A = \begin{bmatrix} 2 & 2 \\ -2 & 7 \end{bmatrix}$. *This matrix has eigenvectors* $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ *and* $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

a) *Find a fundamental matrix,* $\Psi(t)$, *for the system* $\vec{x}' = A\vec{x}$.

b) *Use variation of parameters to solve the non-homogeneous system* $\vec{x}' = A\vec{x} + \begin{bmatrix} e^{6t} \\ 0 \end{bmatrix}$.

c) *If we used method of undetermined coefficients instead, what would be the appropriate guess for the form of the non-homogeneous solution?*

Exercise 532 *Solve* $x_1' = x_2 + t$, $x_2' = x_1 + t$ *with initial conditions* $x_1(0) = 1$, $x_2(0) = 2$, *using diagonalization.*

Exercise 533 *For each of the following vector functions* $\vec{f}(t)$, *find the general solution to the system of differential equations given by*

$$\vec{x}' = \begin{bmatrix} -1 & -4 \\ 2 & 5 \end{bmatrix} \vec{x} + \vec{f}(t)$$

using any of the methods described in this section. Notice the similarities and differences between using these methods for different non-homogeneous parts.

Author(s): Matthew Charnley and Jason Nowell

a) $\vec{f}(t) = \begin{bmatrix} e^{2t} \\ 1 \end{bmatrix}$ b) $\vec{f}(t) = \begin{bmatrix} e^{-t}+2 \\ e^{4t}-1 \end{bmatrix}$ c) $\vec{f}(t) = \begin{bmatrix} e^{3t} \\ t \end{bmatrix}$

d) $\vec{f}(t) = \begin{bmatrix} \sin(3t) \\ 1-\sin(3t) \end{bmatrix}$ e) $\vec{f}(t) = \begin{bmatrix} t+2 \\ e^{-2t} \end{bmatrix}$ f) $\vec{f}(t) = \begin{bmatrix} te^{t} \\ 3 \end{bmatrix}$

Exercise 534 *The variation of parameters method can also be applied to constant coefficient systems. Find the general solution of the system*

$$\vec{x}' = \begin{bmatrix} 4 & -2 \\ 3 & -3 \end{bmatrix}\vec{x} + \begin{bmatrix} e^{2t} \\ e^{t} \end{bmatrix}$$

using

a) *diagonalization* b) *variation of parameters.*

Compare and contrast these methods. You can use undetermined coefficients to check your answer.

Exercise 535 *Find the general solution to the differential equation*

$$\vec{x}' = \begin{bmatrix} -3 & -1 \\ 4 & -3 \end{bmatrix}\vec{x} + \begin{bmatrix} e^{3t}+1 \\ 2 \end{bmatrix}.$$

The best option is undetermined coefficients here because of the eigenvalues of the matrix. Diagonalization can be used, but care will be needed with solving the decoupled system because the coefficients will be complex.

Exercise 536 *Find the general solution to the differential equation*

$$\vec{x}' = \begin{bmatrix} -5 & 16 \\ -1 & 3 \end{bmatrix}\vec{x} + \begin{bmatrix} \cos(2t) \\ \sin(2t)-2\cos(2t) \end{bmatrix}.$$

The best option is undetermined coefficients here because of the eigenvalues of the matrix. We can't actually use diagonalization (try it and see why!).

Exercise 537 *Find the general solution to the differential equation*

$$\vec{x}' = \begin{bmatrix} -2 & -12 \\ 2 & 8 \end{bmatrix}\vec{x} + \begin{bmatrix} e^{2t}+e^{3t} \\ -2e^{2t} \end{bmatrix}.$$

Exercise 538 *Consider the system*

$$\frac{dx}{dt} = x + 2y + 4e^{3t};\ \frac{dy}{dt} = 3x - e^{3t}. \tag{44}$$

a) *Rewrite (44) in the form $\vec{x}' = A\vec{x} + \vec{g}(t)$, where $\vec{x}' = A\vec{x}$ is a homogeneous system, and $\vec{g}(t)$ is a vector-valued function.*

b) *Solve (44) using Method of Undetermined Coefficients.*

Exercise 539

a) *Use variation of parameters to solve the system* $\vec{x}' = \begin{bmatrix} 1 & -4 \\ 4 & -7 \end{bmatrix} \vec{x} + \begin{bmatrix} e^{-3t} \\ 0 \end{bmatrix}$.

b) *What does that solution tell you about how to set up the guess for the method of undetermined coefficients when there is a repeated eigenvalue?*

Exercise 540 *Solve the initial value problem*

$$\vec{x}' = \begin{bmatrix} 5 & -6 \\ 3 & -1 \end{bmatrix} \vec{x} + \begin{bmatrix} t \\ 3 \end{bmatrix} \qquad \vec{x}(0) = \begin{bmatrix} 1 \\ -3 \end{bmatrix}.$$

Exercise 541 *Solve the initial value problem*

$$\vec{x}' = \begin{bmatrix} -4 & 2 \\ -9 & 5 \end{bmatrix} \vec{x} + \begin{bmatrix} e^{3t} \\ e^{t} - 1 \end{bmatrix} \qquad \vec{x}(0) = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

Exercise 542 *Solve the initial value problem*

$$\vec{x}' = \begin{bmatrix} 3 & 2 \\ 0 & 4 \end{bmatrix} \vec{x} + \begin{bmatrix} e^{4t} \\ e^{3t} - t \end{bmatrix} \qquad \vec{x}(0) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Exercise 543 *Take the equation* $\vec{x}' = \begin{bmatrix} \frac{1}{t} & -1 \\ 1 & \frac{1}{t} \end{bmatrix} \vec{x} + \begin{bmatrix} t^2 \\ -t \end{bmatrix}$.

a) *Check that* $\vec{x}_c = c_1 \begin{bmatrix} t \sin t \\ -t \cos t \end{bmatrix} + c_2 \begin{bmatrix} t \cos t \\ t \sin t \end{bmatrix}$ *is the complementary solution.*

b) *Use variation of parameters to find a particular solution.*
