Introduction

This book arose from lecture notes that I began to develop in 2010-2011 for a first course in ordinary differential equations (ODEs). At Calvin College the students in this course are primarily engineers. In our engineering program it is generally the case that the only (formal) linear algebra the students see throughout their undergraduate career is what is presented in the ODE course. Since, in my opinion, the amount of material on linear algebra covered in, e.g., the classical text of Boyce and DiPrima [10], is insufficient if that is all you will see in your academic career, I found it necessary to supplement with notes on linear algebra of my own design. Eventually, it became clear that in order to have a seamless transition between the linear algebra and ODEs, there needed to be one text. This is not a new idea; for example, two recent texts which have a substantive linear algebra component are by Boelkins et al. [7] and Edwards and Penney [16].

Because there is a substantive linear algebra component in this text, I - and more importantly, the students - found it to be much easier later in the text when discussing the solutions of linear systems of ODEs to focus more on the ODE aspects of the problems, and less on the underlying algebraic manipulations. Moreover, I have found that by doing the linear algebra first, it allowed me to more extensively and deeply explore linear systems of ODEs. In particular, it is possible to do much more interesting examples and applications.

The applications presented in this text are labeled “Case Studies”. I chose this moniker because I wanted to convey to the reader that in solving particular problems we were going to do more than simply find a solution; instead, we were going to take time to determine what the solution was telling us about the dynamical behaviour for the given physical system. There are 18 case studies presented herein. Some are classical - e.g., damped mass-spring systems, mixing problems (compartment models) - but several are not typically found in a text such as this. Such examples include a discrete SIR model, a study of the effects on the body of lead ingestion, strongly damped systems (which can be recast as a singular perturbation problem), and a (simple) problem in the mathematics of climate. It is (probably) not possible to present all of these case studies in a one-semester course. On the other hand, the large number (hopefully) allows the instructor to choose a subset which will be of particular interest to his/her class.

The book is formatted as follows. In Chapter 1 we discuss not only the basics of linear algebra that will be needed for solving systems of linear ordinary differential equations, e.g., Gaussian elimination, matrix algebra, and eigenvalues/eigenvectors, but we discuss such foundational material as subspaces, dimension, etc. While the latter material is not necessary to solve ODEs, I find that this is a natural time to introduce students to these more abstract linear algebra concepts. Moreover, since linear algebra is such foundational material for a mathematical understanding of all of the sciences, I feel that it is essential that the students' learn as much as they reasonably can in the short amount of time that is available. It is typically the case that the material in Chapter 1 can be covered in about 15-18 class periods. Primarily because of time constraints, when presenting this material I focus primarily on the case of the vector space $\mathbb{R}^n$. The culminating section in the chapter is that on eigenvalues and eigenvectors. Here I especially emphasize the utility of writing a given vector as a linear combination of the eigenvectors. The closing section considers the large-time behavior associated with three discrete dynamical systems. If the reader and/or instructor wishes to have a supplementary text for this chapter, the book by Hefferon [22] is an excellent companion. Moreover , the PDF can be had for free at `http://joshua.smcvt.edu/linearalgebra/`.

Once the linear algebra has been mastered, we begin the study of ODEs by first solving scalar first-order linear ODEs in Chapter 2. We briefly discuss the general existence/uniqueness theory, as well as the numerical solution. When solving ODEs numerically, we use the MATLAB programs `dfield8.m` and `pplane8.m` developed by J. Polking. These MATLAB programs have accompanying Java applets:

- DFIELD: `http://math.rice.edu/~dfield/dfpp.html`
- PPLANE: `http://math.rice.edu/~dfield/dfpp.html`.

My experience is that these software tools are more than sufficient to numerically solve the problems discussed in this class. We next construct the homogeneous and particular solutions to the linear problem. In this construction we do three things:

(a) derive and write the homogeneous solution formula in such a way that the later notion of a homogeneous solution being thought of as the product of a matrix-valued solution and a constant vector is a natural extension
(b) derive and write the variation-of-parameters solution formula in such a manner that the ideas easily generalize to systems
(c) develop the technique of undetermined coefficients.

The chapter closes with a careful analysis of the one-tank mixing problem under the assumption that the incoming concentration varies periodically in time, and a mathematical finance problem . The idea here is to:

(a) show the students that understanding is not achieved with a solution formula; instead, it is necessary that the formula be written "correctly" so that as much physical information as possible can be gleaned from it
(b) introduce the students to the ideas of amplitude plots and phase plots
(c) set the students up for the later analysis of the periodically forced mass-spring.

As a final note, in many (if not almost all) texts there is typically in this chapter an extensive discussion on nonlinear ODEs. I chose to provide only a cursory treatment of this topic at the end of this book because of:

(a) my desire for my students to understand and focus on linearity and its consequences
(b) the fact that we at Calvin College teach a follow-up course on nonlinear dynamics using the wonderful text by Strogatz [**39**].

In Chapter 3 we study systems of linear ODEs. We start with five physical examples, three of which are mathematically equivalent in that they are modeled by a second-order scalar ODE. We show that n^{th}-order scalar ODEs are equivalent to first-order systems, and thus (hopefully) convince the student that it is acceptable to skip (for the moment) a direct study of these higher-order scalar problems. We almost immediately go the case of the homogeneous problem being constant coefficient, and derive the homogeneous solution via an expansion in terms of the eigenvectors. From a pedagogical perspective I find (and my students seem to agree) this to be a natural way to see how the eigenvalues and eigenvectors of a matrix play a key role in the construction of the homogeneous solution, and in particular how using a particular basis may greatly simplify a given problem. Moreover, I find that this approach serves as an indirect introduction to the notion of Fourier expansions, which is of course used extensively in a successor course on linear partial differential equations. After we construct the homogeneous solutions we discuss the associated phase plane. As for the particular solutions we mimic the discussion of the previous chapter and simply show what few modifications must be made in order for the previous results to be valid for systems. My experience has been that the manner in which things were done in the previous chapter helps the student to see that it is not the case we are learning something entirely new and different, but instead we are just expanding on an already understood concept. The chapter closes with a careful analysis of three problems: a two-tank mixing problem in which the incoming concentration into at one of the tanks is assumed to vary periodically in time, a study of the effect of lead ingestion, and an SIR model associated with zoonotic (animal-to-human) bacterial infections. As in the previous chapter the goal is to not only construct the mathematical solution to the problem, but to also understand how the solution helps us to understand the dynamics of the given physical system.

In Chapter 4 we solve higher-order scalar ODEs. Because all of the theoretical work has already been done in the previous chapter, it is not necessary to spend too much time on this particular task. In particular, there is a relatively short presentation as to how one can use the systems theory to solve the scalar problem. The variation of parameters formula is not re-derived; instead, it is just presented as a special case of the formula for systems. We conclude with a careful study of several problems: the undamped and damped mass-spring systems, a (linear) pendulum driven by a constant torque, a couple mass-spring system, and the vibrations of a beam. The last study introduces the separation of variables technique for solving linear PDEs. Nice illustrative Java applets for the mass-spring problems are:

- Forced and damped oscillations of a spring pendulum:
  `http://www.walter-fendt.de/ph14e/resonance.htm`
- Coupled oscillators:
  `http://www.lon-capa.org/%7emmp/applist/coupled/osc2.htm`.

There are also illustrative movies which are generated by MATLAB.

In Chapter 5 we solve scalar ODEs using the Laplace transform. The focus here is to solve only those problems for which the forcing term is a linear combination of Heaviside functions and delta functions. In my opinion any other type of forcing term can be more easily handled with either the method of undetermined coefficients or variation of parameters. Moreover, we focus on using the Laplace transform as a method to find the particular solution, with the understanding that we can find the homogeneous solution using the ideas and techniques from previous chapters. In order to simplify the calculations, we assume that when finding the particular solution there is zero initial data. Because of the availability of WolframAlpha, we spend little time on partial fraction expansions and the inversion of the Laplace transform. The subsequent case studies are somewhat novel. We start with finding a way to stop the oscillations for an undamped mass-spring system. For our second problem, we study a one-tank mixing problem in which in the incoming concentration varies periodically in time. The injection strategy is modeled as an infinite sum of delta functions. Our last case study involves the analysis of a strongly damped mass-spring problem. We show that this system can be thought of as a singular perturbation problem which is (formally) mathematically equivalent to a one-tank mixing problem. We finish the discussion of the Laplace transform with the engineering applications of the transfer function, the manner in which the poles of the transfer function effect the dynamics of the homogeneous solution. We show that the convolution integral leads to a variation-of-parameters formula for the particular solution.

In Chapter 6 we cover topics which are not infrequently discussed if time permits: separation of variables, phase line analysis, and series solutions. Each topic is only briefly touched upon, but enough material is presented herein for the student to get a good idea of what each one is about. For the latter two topics I present case studies which could lead to a more detailed examination of the topic (using outside resources) if the student and/or instructor wishes.

Almost every section concludes with a set of homework problems. Moreover, there is a section at the end of each of Chapter 1, Chapter 3, Chapter 4, and Chapter 5 which is labeled Group Projects. The problems contained in these sections are more challenging, and I find it to be the case that the students have a better chance of understanding and solving them if they work together in groups of 3-4 people. My experience is that the students truly enjoy working on these problems, and they very much appreciate working collaboratively. I typically assign 1-2 of these types of problems per semester.

As of the current edition relatively few of the homework problems have attached to them a solution. My expectation is that many, if not most, students will find this lack of solved problems troubling. Two relatively cheap (potentially supplemental) texts which address this issue are Lipschutz and Lipson [**28**] for the linear algebra material and Bronson and Costa [**11**] for the ODE material. Of course, other books, e.g., [**6**, **13**, **19**, **34**], can be found simply by going to the library and looking there through the (perhaps) dozens of appropriate books.

Throughout this text we expect the students to use a CAS to do some of the intermediate calculations. Herein we focus upon WolframAlpha (`http://www.wolframalpha.com/`). There are several advantages to using this particular CAS:

(a) it is not necessary to learn a programming language to use it
(b) the commands are intuitive
(c) it is easily accessible
(d) it is free (as of June, 2014).

I appreciate that the interested reader and/or instructor can do much more with Mathematica, Maple, Sage, etc. However, there is currently no universal agreement as to which package is best to use (even within my department!), and I do not want to limit this text to a particular system. Moreover, my goal here is to focus more on using the software to solve a given problem, and not on the programming necessary to use the particular CAS. My expectation is that interested students who have some experience with a particular CAS will quickly learn how to do what they want to do with it.

In this text we do not use this software to completely solve a given problem, as it is important that the student thoroughly understand what intermediate calculations are needed in order to solve the problem. The idea here is that the CAS can be used to remove some of the computational burden associated with solving a problem. A screenshot is provided in the text for most of the calculations, so it should be easy for the student to replicate. In addition, there is a brief section at the end of the text which shows how one can use MATLAB to perform many of the intermediate calculations. The particular scripts are provided on my web page at http://www.calvin.edu/~tmk5/courses/m231/S14/.

In the ODE portion of this text we attempt to emphasize the idea that the interesting thing is not necessarily the mathematical solution of a given mathematical problem, but what it is that the solution tells you about the physical problem being modeled. The (extremely) easy-to-use CAS generally does a reasonable job of solving a given mathematical equation, but it is not quite as helpful when interpreting a solution.

The electronic version of this book is embedded with hyperlinks (both internal and external), and they are marked in blue text. It is my hope that these links make it easier to navigate the book; in particular, it should be the case that it is easier (and quicker than a paper version!) for the reader to reference previous results, e.g., to recall a result on page 69 while reading page 113. The book does include a minimal index. It primarily gives the first page at which particular term is mentioned. In particular, it does provide the page for which each term is first defined. Since it is expected that this book will be primarily used in an electronic format, this potential drawback is easily overcome via a "find" command.

I am indebted to Kate Ardinger, Tom Jager, Jeff Humpherys, Michael Kapitula, Keith Promislow, Thomas Scofield, Matt Walhout, and anonymous reviewers for discussions about, and a careful reading of, this manuscript. The implementation of their suggestions and comments greatly improved the text.

For the glory of the most high God alone,
And for my neighbour to learn from.
— *J.S. Bach* —

Chapter 1
Essentials of Linear Algebra

Mathematics is the art of reducing any problem to linear algebra.
- William Stein

. . . mathematics is a collection of examples; a theorem is a statement about a collection of examples and the purpose of proving theorems is to classify and explain the examples . . .
- John Conway

The average college student knows how to solve two equations in two unknowns in an elementary way: the method of substitution. For example, consider the system of equations

$$2x + y = 6, \quad 2x + 4y = 5.$$

Solving the first equation for y gives $y = 6 - 2x$, and substituting this expression into the second equation yields

$$2x + 4(6 - 2x) = 5 \quad \Rightarrow \quad x = \frac{19}{6}.$$

Substitution into either of the equations gives the value of y; namely, $y = -1/3$. For systems of three or more equations this algorithm is algebraically unwieldy. Furthermore, it is inefficient, as it is often the case not very clear as to which variable(s) should be substituted into which equation(s). Thus, at the very least, we should develop an efficient algorithm for solving large systems of equations. Perhaps more troubling (at least to the mathematician!) is the fact that the method of substitution does not yield any insight into the structure of the solution set. An analysis and understanding of this structure is the topic of linear algebra. As we will see, not only will we gain a much better understanding of how to solve linear algebraic systems, but by considering the problem more abstractly we will better understand how to solve linear systems of ordinary differential equations (ODEs).

This chapter is organized in the following manner. We begin our discussion of linear systems of equations by developing an efficient solution algorithm: Gaussian elimination. We then consider the problem using matrices and vectors, and spend considerable time and energy trying to understand the solution structure via these objects. In particular, we show that that the solution is composed of two pieces. One piece intrinsically associated with the matrix alone, and the other piece reflects an interaction between the matrix and nonhomogeneous term. We conclude the chapter by looking at special vectors associated with square matrices: the eigenvectors. These vectors have the special algebraic property that the matrix multiplied by the eigenvector is simply a scalar multiple of the eigenvector (this scalar is known as the associated eigenvalue). As we will see, the eigenvalues and eigenvectors are the key objects associated with a matrix that allow us to easily and explicitly write down and understand the solution to a linear dynamical systems (both discrete and continuous).

1.1 Solving linear systems

1.1.1 *Notation and terminology*

A *linear equation* in n variables is an algebraic equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b. \tag{1.1.1}$$

The (possibly complex-valued) numbers $a_1, a_2, \ldots, a_n$ are the *coefficients*, and the unknowns to be solved for are the *variables* $x_1, \ldots, x_n$. The variables are also sometimes called *unknowns*. An example in two variables is

$$2x_1 - 5x_2 = 7,$$

and an example in three variables is

$$x_1 - 3x_2 + 9x_3 = -2.$$

A *system of linear equations* is a collection of m linear equations (1.1.1), and can be written as

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{1.1.2}$$

The coefficient a_{jk} is associated with the variable x_k in the j^{th} equation. An example of two equations in three variables is

$$\begin{aligned} x_1 - 4x_2 \qquad &= 6 \\ 3x_1 + 2x_2 - 5x_3 &= 2. \end{aligned} \tag{1.1.3}$$

> Until we get to our discussion of eigenvalues and eigenvectors in Chapter 1.13, we will assume that the coefficients and variables are real numbers, i.e., $a_{jk}, x_j \in \mathbb{R}$. This is done solely for the sake of pedagogy and exposition. It cannot be stressed too much, however, that *everything* we do preceding Chapter 1.13 still works even if we remove this restriction, and we allow these numbers to be complex-valued.

When there is a large number of equations and/or variables, it is awkward to write down a linear system in the form of (1.1.2). It is more convenient instead to use a *matrix* formulation. A matrix is a rectangular array of numbers with m rows and n columns, and such a matrix is said to be an $m \times n$ (read "m by n") matrix. If $m = n$, the matrix is said to be a *square matrix*. A matrix will be denoted as $\boldsymbol{A}$, and we will say $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. The *coefficient matrix* for the linear system (1.1.2) is given by

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \tag{1.1.4}$$

and the coefficient a_{jk}, which is associated with the variable x_k in the j^{th} equation, is in the j^{th} row and k^{th} column. For example, the coefficient matrix for the system (1.1.3) is given by

$$\boldsymbol{A} = \begin{pmatrix} 1 & -4 & 0 \\ 3 & 2 & -5 \end{pmatrix} \in \mathbb{R}^{2\times 3},$$

with

$$a_{11} = 1,\ a_{12} = -4,\ a_{13} = 0,\ a_{21} = 3,\ a_{22} = 2,\ a_{23} = -5.$$

A *vector*, say $\boldsymbol{v}$, is an $m \times 1$ matrix, i.e., a matrix with only one column. A vector is sometimes called a *column vector* or *m-vector*, and we write $\boldsymbol{v} \in \mathbb{R}^{m\times 1} := \mathbb{R}^m$. The variables in the system (1.1.2) will be written as the vector

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix},$$

and the variables on the right-hand side will be written as the vector

$$\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

In conclusion, for the system (1.1.2) there are three matrix-valued quantities: $\boldsymbol{A}$, $\boldsymbol{x}$ and $\boldsymbol{b}$. We will represent the linear system (1.1.2)

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \tag{1.1.5}$$

We will later see what it means to multiply a matrix and a vector. The linear system is said to be *homogeneous* if $\boldsymbol{b} = \boldsymbol{0}$, i.e., $b_j = 0$ for $j = 1, \dots, m$; otherwise, the system is said to be *nonhomogeneous*.

1.1.2 Solutions of linear systems

A *solution* to the linear system (1.1.5) (or equivalently, (1.1.2)) is a vector $\boldsymbol{x}$ which satisfies all m equations simultaneously. For example, consider the linear system of three equations in three unknowns for which

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & -1 \\ 3 & 1 & 0 \\ 1 & -1 & -1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 0 \\ 1 \\ -4 \end{pmatrix}, \tag{1.1.6}$$

i.e.,

$$x_1 - x_3 = 0, \quad 3x_1 + x_2 = 1, \quad x_1 - x_2 - x_3 = -4.$$

It is not difficult to check that a solution is given by

$$\boldsymbol{x} = \begin{pmatrix} -1 \\ 4 \\ -1 \end{pmatrix} \quad \Rightarrow \quad x_1 = -1,\ x_2 = 4,\ x_3 = -1.$$

A system of linear equations with at least one solution is said to be *consistent*; otherwise, it is *inconsistent*.

How many solutions does a linear system have? Consider the system given by

$$2x_1 - x_2 = -2, \quad -x_1 + 3x_3 = 11.$$

The first equation represents a line in the x_1x_2-plane with slope 2, and the second equation represents a line with slope 1/3. Since lines with different slopes intersect at a unique point, there is a unique solution to this system, and it is consistent. It is not difficult to check that the solution is given by $(x_1, x_2) = (1, 4)$. Next consider the system given by

$$2x_1 - x_2 = -2, \quad -4x_1 + 2x_3 = 8.$$

Each equation represents a line with slope 2, so that the lines are parallel. Consequently, the lines are either identically the same, so that there are an infinite number of solutions, or they intersect at no point, so that the system is inconsistent. Since the second equation is a multiple of the first equation, the system is consistent. On the other hand, the system

$$2x_1 - x_2 = -2, \quad -4x_1 + 2x_3 = 7$$

is inconsistent, as the second equation is no longer a scalar multiple of the first equation. See Figure 1.1 for graphical representations of these three cases.

We see that a linear system with two equations and two unknowns is either consistent with one or an infinite number of solutions, or is inconsistent. It is not difficult to show that this fact holds for linear systems with three unknowns. Each linear equation in the system represents a plane in $x_1x_2x_3$-space. Given any two planes, we know that they are either parallel, or intersect along a line. Thus, if the system has two equations, then it will either be consistent with an infinite number of solutions, or inconsistent. Suppose that the system with two equations is consistent, and add a third linear equation. Further suppose that the original two planes intersect along a line. This new plane is either parallel to the line, or intersects it at precisely one point. If the original two planes are the same, then the new plane is either parallel to both, or intersects it along a line. In conclusion, for a system of equations with three variables there is either a unique solution, an infinite number of solutions, or no solution.

This argument can be generalized to show the following result:

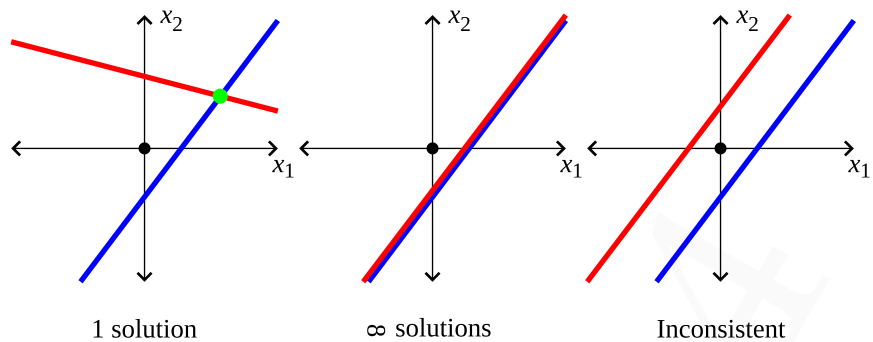


Fig. 1.1 (color online) A graphical depiction of the three possibilities for linear systems of two equations in two unknowns. The left panel shows the case when the corresponding lines are not parallel, and the other two panels show the cases when the lines are parallel.

Theorem 1.1.1. *If the linear system (1.1.2) is consistent, then there is either a unique solution, or an infinite number of solutions.*

▷ *Remark 1.1.2.* Theorem 1.1.1 does not hold for nonlinear systems. For example, the nonlinear system

$$x_1^2 + x_2^2 = 2, \quad x_1 + x_2 = 0$$

is consistent, and has the two solutions $(-1, 1)$, $(1, -1)$.

It is often the case that if a linear system is consistent, then more cannot be said about the number of solutions without directly solving the system. However, in the argument leading up to Theorem 1.1.1 we did see that for a system of two equations in three unknowns that if the system was consistent, then there were necessarily an infinite number of solutions. This result holds in general:

Corollary 1.1.3. *Suppose that the linear system is such that $m < n$, i.e., there are fewer equations than unknowns (the system is underdetermined). If the system is consistent, then there are an infinite number of solutions.*

1.1.3 Solving by Gaussian elimination

We now need to understand how to systematically solve the linear system (1.1.5),

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

While the method of substitution works fine for two equations in two unknowns, it quickly breaks down as a practical method when there are three or more variables involved in the system. We need to come up with something else.

The simplest linear system to solve for two equations in two unknowns is

The identity matrix, $\boldsymbol{I}_n$, is a square matrix with ones of the diagonal, and zeros everywhere else. The subscript refers to the size of the matrix.

$$x_1 = b_1, \quad x_2 = b_2.$$

The coefficient matrix is

$$\boldsymbol{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2},$$

which is known as the *identity matrix*. The unique solution to this system is $\boldsymbol{x} = \boldsymbol{b}$. The simplest linear system to solve for three equations in three unknowns is

$$x_1 = b_1, \quad x_2 = b_2, \quad x_3 = b_3.$$

The coefficient matrix is now

$$\boldsymbol{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3},$$

which is still known as the identity matrix. The unique solution to again system is $\boldsymbol{x} = \boldsymbol{b}$. Continuing in this fashion, the simplest linear system for n equations in n unknowns to solve is

$$x_1 = b_1,\ x_2 = b_2,\ x_3 = b_3, \dots, x_n = b_n.$$

The coefficient matrix associated with this system is $\boldsymbol{I}_n$, and the solution is $\boldsymbol{x} = \boldsymbol{b}$.

Suppose that the number of equations is not equal to the number of unknowns. For example, a particularly simple system to solve is given by

$$x_1 - 3x_3 + 4x_4 = 2, \quad x_2 + x_3 - 6x_4 = 5. \tag{1.1.7}$$

The coefficient matrix for this system is

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & -3 & 4 \\ 0 & 1 & 1 & -6 \end{pmatrix} \in \mathbb{R}^{2\times 4}.$$

Solving the system for the first two variables in terms of the latter two yields

$$x_1 = 3x_3 - 4x_4 + 2, \quad x_2 = -x_3 + 6x_4 + 5.$$

Upon setting $x_3 = s$ and $x_4 = t$, where the dummy variables $s, t \in \mathbb{R}$ are arbitrary, we see the solution to this system is

$$x_1 = 2 + 3s - t,\ x_2 = 5 - s + 6t,\ x_3 = s,\ x_4 = t \quad \Rightarrow \quad \boldsymbol{x} = \begin{pmatrix} 2 + 3s - 4t \\ 5 - s + 6t \\ s \\ t \end{pmatrix}, \quad s, t \in \mathbb{R}.$$

Since s and t are arbitrary, there are an infinite number of solutions. This was expected, for as we saw in Corollary 1.1.3 consistent underdetermined systems will have an infinite number of solutions.

The coefficient matrices for the problems considered so far share a common feature, which is detailed below:

RREF

Definition 1.1.4. A matrix is said to be in *row reduced echelon form* (RREF) if

(a) all nonzero rows are above any zero row
(b) the first nonzero entry in a row (the *leading entry*) is a one
(c) every other entry in a column with a leading one is zero.

Those columns with a leading entry are known as *pivot columns*, and the leading entries are called *pivot positions*.

◁ *Example* 1.1.5. Consider the matrix in RREF given by

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & -3 & 0 & 7 \\ 0 & 1 & -3 & 0 & 2 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{4\times 5}$$

The first, second, and fourth columns are the pivot columns, and the pivot positions are the first entry in the first row, the second entry in the second row, and the fourth entry in the third row.

If a coefficient matrix is in RREF, then the linear system is particulary easy to solve. Thus, our goal is to take a given linear system with its attendant coefficient matrix, and then perform allowable algebraic operations so that the new system has a coefficient matrix which is in RREF. The allowable algebraic operations for solving a linear system are:

(a) multiply any equation by a constant
(b) add/subtract equations
(c) switch the ordering of equations.

As a rule-of-thumb, when putting an augmented matrix into RREF, the idea is to place 1's on the diagonal, and 0's everywhere else (as much as possible).

In order to do these operations most efficiently using matrices, it is best to work with the *augmented matrix* associated with the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$; namely, the matrix $(\boldsymbol{A}|\boldsymbol{b})$. The augmented matrix is formed by adding a column, namely the vector $\boldsymbol{b}$, to the coefficient matrix. For example, for the linear system associated with (1.1.6) the augmented matrix is given by

$$(\boldsymbol{A}|\boldsymbol{b}) = \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 3 & 1 & 0 & 1 \\ 1 & -1 & -1 & -4 \end{array}\right), \tag{1.1.8}$$

and the augmented matrix for the linear system (1.1.7) is

$$(\boldsymbol{A}|\boldsymbol{b}) = \left(\begin{array}{cccc|c} 1 & 0 & -3 & 4 & 2 \\ 0 & 1 & 1 & -6 & 5 \end{array}\right)$$

The allowable operations on the individual equations in the linear system correspond to operations on the *rows* of the augmented matrix. In particular, when doing *Gaussian elimination* on an augmented matrix in order to put it into RREF, we are allowed to:

(a) multiply any row by a constant
(b) add/subtract rows
(c) switch the ordering of the rows.

Once we have performed Gaussian elimination on an augmented matrix in order to put it into RREF, we can easily solve the resultant system.

◁ *Example 1.1.6.* Consider the linear system associated with the augmented matrix in (1.1.8). We will henceforth let $\boldsymbol{\rho}_j$ denote the j^{th} row of a matrix. The operation "$a\rho_j + b\rho_k$" will be taken to mean multiply the j^{th} row by a, multiply the k^{th} row by b, add the two resultant rows together, and replace the k^{th} row with this sum. With this notation in mind, performing Gaussian elimination yields

$$(\boldsymbol{A}|\boldsymbol{b}) \xrightarrow{-3\boldsymbol{\rho}_1+\boldsymbol{\rho}_2} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & 1 \\ 1 & -1 & -1 & -4 \end{array}\right) \xrightarrow{-\boldsymbol{\rho}_1+\boldsymbol{\rho}_3} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & -1 & 0 & -4 \end{array}\right) \xrightarrow{\boldsymbol{\rho}_2+\boldsymbol{\rho}_3} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 3 & -3 \end{array}\right)$$

$$\xrightarrow{(1/3)\boldsymbol{\rho}_3} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right) \xrightarrow{-3\boldsymbol{\rho}_3+\boldsymbol{\rho}_2} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -1 \end{array}\right) \xrightarrow{\boldsymbol{\rho}_3+\boldsymbol{\rho}_1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -1 \end{array}\right).$$

The new linear system is

$$x_1 = -1,\ x_2 = 4,\ x_3 = -1 \quad \Rightarrow \quad \boldsymbol{x} = \begin{pmatrix} -1 \\ 4 \\ -1 \end{pmatrix},$$

which is also immediately seen to be the solution.

◁ *Example 1.1.7.* Consider the linear system

$$\begin{aligned} x_1 - 2x_2 - x_3 &= 0 \\ 3x_1 + x_2 + 4x_3 &= 7 \\ 2x_1 + 3x_2 + 5x_3 &= 7. \end{aligned}$$

Performing Gaussian elimination on the augmented matrix yields

$$\left(\begin{array}{ccc|c} 1 & -2 & -1 & 0 \\ 3 & 1 & 4 & 7 \\ 2 & 3 & 5 & 7 \end{array}\right) \xrightarrow{-3\boldsymbol{\rho}_1+\boldsymbol{\rho}_2} \left(\begin{array}{ccc|c} 1 & -2 & -1 & 0 \\ 0 & 7 & 7 & 7 \\ 2 & 3 & 5 & 7 \end{array}\right) \xrightarrow{-2\boldsymbol{\rho}_1+\boldsymbol{\rho}_3} \left(\begin{array}{ccc|c} 1 & -2 & -1 & 0 \\ 0 & 7 & 7 & 7 \\ 0 & 7 & 7 & 7 \end{array}\right) \xrightarrow{-\boldsymbol{\rho}_2+\boldsymbol{\rho}_3} \left(\begin{array}{ccc|c} 1 & -2 & -1 & 0 \\ 0 & 7 & 7 & 7 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\xrightarrow{(1/7)\boldsymbol{\rho}_2} \left(\begin{array}{ccc|c} 1 & -2 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{2\boldsymbol{\rho}_2+\boldsymbol{\rho}_1} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

The new linear system to be solved is given by

$$x_1 + x_3 = 2, \quad x_2 + x_3 = 1, \quad 0x_1 + 0x_2 + 0x_3 = 0.$$

Ignoring the last equation, this is a system of two equations with three unknowns; consequently, since the system is consistent it must be the case that there are an infinite number of solutions. The variables x_1 and x_2 are associated with leading entries in the RREF form of the augmented matrix. As for the variable x_3, which is not associated with a leading entry:

Free variable

Definition 1.1.8. Any variable of a RREF matrix which is not associated with a leading entry is a *free variable*. In other words, free variables are associated with those columns in the RREF matrix which are not pivot columns.

Since x_3 is a free variable, it can be arbitrarily chosen. Upon setting $x_3 = t$, where $t \in \mathbb{R}$, the other variables are

$$x_1 = 2 - t, \quad x_2 = 1 - t.$$

The solution is then

$$\boldsymbol{x} = \begin{pmatrix} 2-t \\ 1-t \\ t \end{pmatrix}, \quad t \in \mathbb{R}.$$

◁ *Example 1.1.9.* Consider a linear system which is a variant of the one given above; namely,

$$\begin{aligned} x_1 - 2x_2 - x_3 &= 0 \\ 3x_1 + x_2 + 4x_3 &= 7 \\ 2x_1 + 3x_2 + 5x_3 &= 8. \end{aligned}$$

Upon doing Gaussian elimination of the augmented matrix we see that

$$\left(\begin{array}{rrr|r} 1 & -2 & -1 & 0 \\ 3 & 1 & 4 & 7 \\ 2 & 3 & 5 & 8 \end{array}\right) \longrightarrow \left(\begin{array}{rrr|r} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

The new linear system to be solved is

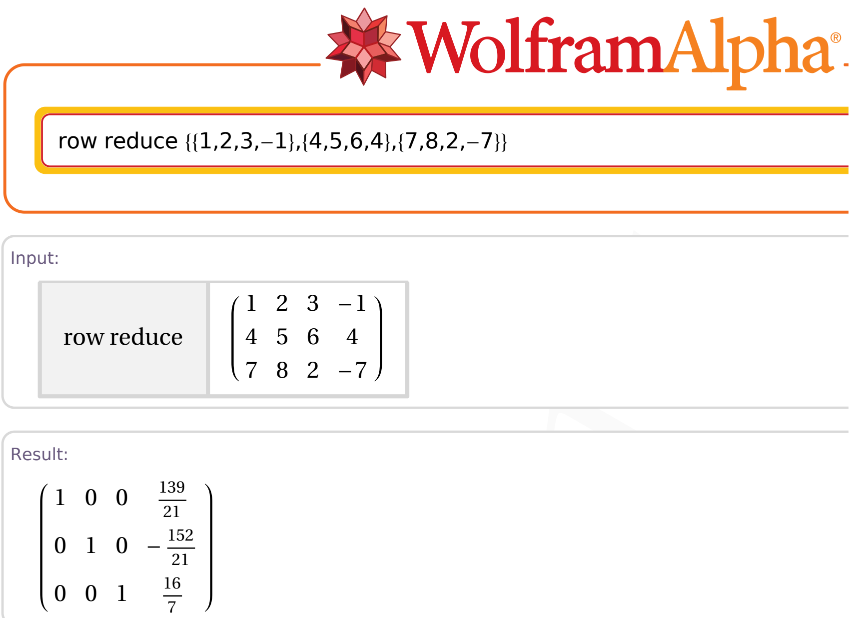
$$x_1 + x_3 = 2, \quad x_2 + x_3 = 1, \quad 0x_1 + 0x_2 + 0x_3 = 1.$$

Since the last equation clearly does not have a solution, the system is inconsistent.

◁ *Example 1.1.10.* Consider a linear system for which the coefficient matrix and nonhomogeneous term are

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 2 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} -1 \\ 4 \\ -7 \end{pmatrix}.$$

We will use WolframAlpha to put the augmented matrix into RREF. It is straightforward to enter a matrix in this CAS. The full matrix is surrounded by curly brackets. Each individual row is also surrounded by curly brackets, and the individual entries in a row are separated by commas. Each row is also separated by a comma. We have:



$$\text{row reduce} \begin{pmatrix} 1 & 2 & 3 & -1 \\ 4 & 5 & 6 & 4 \\ 7 & 8 & 2 & -7 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & \frac{139}{21} \\ 0 & 1 & 0 & -\frac{152}{21} \\ 0 & 0 & 1 & \frac{16}{7} \end{pmatrix}$$

The solution is the last column,

$$\boldsymbol{x} = \frac{1}{21}\begin{pmatrix} 139 \\ -152 \\ 48 \end{pmatrix} \sim \begin{pmatrix} 6.62 \\ -7.24 \\ 2.29 \end{pmatrix}.$$

Exercises

Exercise 1.1.1. Solve each system of equations, or explain why no solution exists.

(a) $x_1 + 2x_2 = 4,\ -2x_1 + 3x_2 = -1$
(b) $x_1 + 2x_2 = 4,\ x_1 + 2x_2 = -1$
(c) $x_1 + 2x_2 = 4,\ 4x_1 + 8x_2 = 15$

Exercise 1.1.2. Each of the below linear systems is represented by an augmented matrix in RREF. If the system is consistent, express the solution in vector form.

(a) $\left(\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 5 \end{array}\right)$

(b) $\left(\begin{array}{ccc|c} 1 & 0 & 0 & -4 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 0 \end{array}\right)$

(c) $\left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & -2 \end{array}\right)$

(d) $\left(\begin{array}{cccc|c} 1 & 0 & 0 & 3 & -1 \\ 0 & 1 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$

Exercise 1.1.3. Determine all value(s) of r which make each augmented matrix correspond to a consistent linear system. For each such r, express the solution to the corresponding linear system in vector form.

(a) $\left(\begin{array}{rr|r} 1 & 4 & -3 \\ -2 & -8 & r \end{array}\right)$

(b) $\left(\begin{array}{rr|r} 1 & 4 & -3 \\ 2 & r & -6 \end{array}\right)$

(c) $\left(\begin{array}{rr|r} 1 & 4 & -3 \\ -3 & r & -9 \end{array}\right)$

(d) $\left(\begin{array}{rr|r} 1 & r & -3 \\ -3 & r & 8 \end{array}\right)$

Exercise 1.1.4. The augmented matrix for a linear system is given by

$$\left(\begin{array}{rrr|r} 1 & 1 & 3 & 2 \\ 1 & 2 & 4 & 3 \\ 1 & 3 & a & b \end{array}\right).$$

(a) For what value(s) of a and b will the system have infinitely many solutions?
(b) For what value(s) of a and b will the system be inconsistent?

Exercise 1.1.5. Solve each linear system, and express the solution in vector form.

(a) $3x_1 + 2x_2 = 16, \quad -2x_1 + 3x_2 = 11$
(b) $3x_1 + 2x_2 - x_3 = -2, \quad -3x_1 - x_2 + x_3 = 5, \quad 3x_1 + 2x_2 + x_3 = 2$
(c) $2x_1 + x_2 = -1, \quad x_1 - x_3 = -2, \quad -x_1 + 3x_2 + 7x_3 = 11$
(d) $x_1 + x_2 - x_3 = 0, \quad 2x_1 - 3x_2 + 5x_3 = 0, \quad 4x_1 - x_2 + 3x_3 = 0$
(e) $x_2 + x_3 - x_4 = 0, \quad x_1 + x_2 + x_3 + x_4 = 6$
$2x_1 + 4x_2 + x_3 - 2x_4 = -1, \quad 3x_1 + x_2 - 2x_3 + 2x_4 = 3$

Exercise 1.1.6. If the coefficient matrix satisfies $\boldsymbol{A} \in \mathbb{R}^{9\times 6}$, and if the RREF of the augmented matrix $(\boldsymbol{A}|\boldsymbol{b})$ has three zero rows, is the solution unique? Why, or why not?

Exercise 1.1.7. If the coefficient matrix satisfies $\boldsymbol{A} \in \mathbb{R}^{5\times 7}$, and if the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent, is the solution unique? Why, or why not?

Exercise 1.1.8. Determine if each of the following statements is true or false. Provide an explanation for your answer.

(a) A system of four linear equations in three unknowns can have exactly five solutions.
(b) If a system has a free variable, then there will be an infinite number of solutions.
(c) If a system is consistent, then there is a free variable.
(d) If the RREF of the augmented matrix has four zero rows, and if the system is consistent, then there will be an infinite number of solutions.
(e) If the RREF of the augmented matrix has no zero rows, then the system is consistent.

Exercise 1.1.9. Find a quadratic polynomial $p(t) = a_0 + a_1 t + a_2 t^2$ which passes through the points $(-2, 12), (1, 6), (2, 18)$. *Hint*: $p(1) = 6$ implies that $a_0 + a_1 + a_2 = 6$.

Exercise 1.1.10. Find a cubic polynomial $p(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$ which passes through the points $(-1, -3), (0, 1), (1, 3), (2, 17)$.

1.2 Linear combinations of vectors and matrix/vector multiplication

Now that we have an efficient algorithm to solve the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, we need to next understand what it means from a geometric perspective to solve the system. To begin with, if the system is consistent, how does the vector $\boldsymbol{b}$ relate to the coefficients of the coefficient matrix $\boldsymbol{A}$? In order to answer this question, we first need to make sense of the expression $\boldsymbol{A}\boldsymbol{x}$ (matrix/vector multiplication).

1.2.1 Linear combinations of vectors

We will define the addition/subtraction of two n-vectors to be exactly what is expected, and the same will hold true for the multiplication of a vector by a scalar; namely,

$$\boldsymbol{x} \pm \boldsymbol{y} = \begin{pmatrix} x_1 \pm y_1 \\ x_2 \pm y_2 \\ \vdots \\ x_n \pm y_n \end{pmatrix}, \quad c\boldsymbol{x} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}.$$

Vector addition and subtraction are done component-by-component, and scalar multiplication of a vector means that each component of the vector is multiplied by the scalar. For example,

$$\begin{pmatrix} -2 \\ 5 \end{pmatrix} + \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \quad 3\begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 6 \\ -9 \end{pmatrix}.$$

These are *linear operations*. Combining these two operations, we have more generally:

Linear combination

Definition 1.2.1. A *linear combination* of the n-vectors $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ is given by the vector $\boldsymbol{b}$, where

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k = \sum_{j=1}^{k} x_j\boldsymbol{a}_j.$$

The scalars $x_1, \dots, x_k$ are known as *weights*.

With this notion of linear combinations of vectors, we can rewrite linear systems of equations in vector notation. For example, consider the linear system

$$\begin{aligned} x_1 - x_2 + x_3 &= -1 \\ 3x_1 + 2x_2 + 8x_3 &= 7 \\ -2x_1 - 4x_2 - 8x_3 &= -10. \end{aligned} \tag{1.2.1}$$

Upon using the fact that two vectors are equal if and only if all of their coefficients are equal, we can write (1.2.1) in vector form as

$$\begin{pmatrix} x_1 - x_2 + x_3 \\ 3x_1 + 2x_2 + 8x_3 \\ -2x_1 - 4x_2 - 8x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 7 \\ -10 \end{pmatrix}.$$

Using linearity we can write the vector on the left-hand side as

$$\begin{pmatrix} x_1 - x_2 + x_3 \\ 3x_1 + 2x_2 + 8x_3 \\ -2x_1 - 4x_2 - 8x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 2 \\ -4 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 8 \\ -8 \end{pmatrix},$$

so the system (1.2.1) is equivalent to

$$x_1 \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 2 \\ -4 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 8 \\ -8 \end{pmatrix} = \begin{pmatrix} -1 \\ 7 \\ -10 \end{pmatrix}.$$

After setting

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} -1 \\ 2 \\ -4 \end{pmatrix}, \quad \boldsymbol{a}_3 = \begin{pmatrix} 1 \\ 8 \\ -8 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} -1 \\ 7 \\ -10 \end{pmatrix},$$

the linear system can then be rewritten as the linear combination of vectors

$$x_1 \boldsymbol{a}_1 + x_2 \boldsymbol{a}_2 + x_3 \boldsymbol{a}_3 = \boldsymbol{b}. \tag{1.2.2}$$

In conclusion, asking for solutions to the linear system (1.2.1) can instead be thought of as asking if the vector $\boldsymbol{b}$ is a linear combination of the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$. It can be checked that after Gaussian elimination

$$\left(\begin{array}{rrr|r} 1 & -1 & 1 & -1 \\ 3 & 2 & 8 & 7 \\ -2 & -4 & -8 & -10 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{rrr|r} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

so that the solution to the linear system (1.2.1) is given by

$$x_1 = 1 - 2t,\ x_2 = 2 - t,\ x_3 = t; \quad t \in \mathbb{R}. \tag{1.2.3}$$

The vector $\boldsymbol{b}$ is a linear combination of the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$, and the weights are given in (1.2.3),

$$\boldsymbol{b} = (1 - 2t)\boldsymbol{a}_1 + (2 - t)\boldsymbol{a}_2 + t\boldsymbol{a}_3, \quad t \in \mathbb{R}.$$

1.2.2 Matrix/vector multiplication

With this observation in mind, we now define the multiplication of a matrix and a vector so that the resultant corresponds to a linear system. For the linear system of (1.2.1) let $\boldsymbol{A}$ be the coefficient matrix,

$$\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \boldsymbol{a}_3) \in \mathbb{R}^{3\times 3}.$$

Here each column of $\boldsymbol{A}$ is thought of as a vector. If for

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

we define

$$\boldsymbol{A}\boldsymbol{x} := x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3,$$

then by using (1.2.2) we have that the linear system is given by

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{1.2.4}$$

(compare with (1.1.5)). In other words, by writing the linear system in the form of (1.2.4) we really mean the linear combinations of (1.2.2), which in turn is equivalent to the original system (1.2.1).

Matrix/vector multiplication

Definition 1.2.2. Suppose that $\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2 \cdots \boldsymbol{a}_n)$, where each vector $\boldsymbol{a}_j \in \mathbb{R}^m$ is an m-vector. For $\boldsymbol{x} \in \mathbb{R}^n$ we define *matrix/vector multiplication* as

$$\boldsymbol{A}\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \sum_{j=1}^{n} x_j\boldsymbol{a}_j.$$

Note that $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{x} \in \mathbb{R}^n = \mathbb{R}^{n\times 1}$, so by definition

$$\underbrace{\boldsymbol{A}}_{\mathbb{R}^{m\times n}}\underbrace{\boldsymbol{x}}_{\mathbb{R}^{n\times 1}} = \underbrace{\boldsymbol{b}}_{\mathbb{R}^{m\times 1}}.$$

In order for a matrix/vector multiplication to make sense, the number of columns in the matrix $\boldsymbol{A}$ must the be same as the number of entries in the vector $\boldsymbol{x}$. The product will be a vector in which the number of entries is equal to the number of rows in $\boldsymbol{A}$.

◁ *Example 1.2.3.* We have

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} -3 \\ 5 \end{pmatrix} = -3\begin{pmatrix} 1 \\ 3 \end{pmatrix} + 5\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 7 \\ 11 \end{pmatrix},$$

and

$$\begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 4 \end{pmatrix} + 3\begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 15 \\ 20 \end{pmatrix}.$$

Note that in the first example a 2×2 matrix multiplied a 2×1 matrix in order to get a 2×1 matrix, whereas in the second example a 2×3 matrix multiplied a 3×1 matrix in order to get a 2×1 matrix.

The multiplication of a matrix and a vector is a *linear operation*, as it satisfies the property that the product of a matrix with a linear combination of vectors is the same thing as first taking the individual matrix/vector products, and then taking the appropriate linear combination of the resultant two vectors:

Lemma 1.2.4. *If $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ with $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$, then*

$$\boldsymbol{A}(c\boldsymbol{x} + d\boldsymbol{y}) = c\boldsymbol{A}\boldsymbol{x} + d\boldsymbol{A}\boldsymbol{y}.$$

Proof. Writing $\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2 \cdots \boldsymbol{a}_n)$, and using the fact that

$$cx+dy=\begin{pmatrix}cx_1+dy_1\\cx_2+dy_2\\\vdots\\cx_n+dy_n\end{pmatrix},$$

we have

$$\begin{aligned}A(cx+dy)&=(cx_1+dy_1)a_1+(cx_2+dy_2)a_2+\cdots+(cx_n+dy_n)a_n\\&=[cx_1a_1+cx_2a_2+\cdots+cx_na_n]+[dy_1a_1+dy_2a_2+\cdots dy_na_n]\\&=c[x_1a_1+x_2a_2+\cdots+x_na_n]+d[y_1a_1+y_2a_2+\cdots+y_na_n]\\&=cAx+dAy.\end{aligned}$$

□

▷ *Remark 1.2.5.* We are already familiar with *linear operators*, which are simply operators which satisfy the linearity property of Lemma 1.2.4, in other contexts. For example, if D represents differentiation, i.e., $\mathrm{D}[f(t)]=f'(t)$, then we know from Calculus I that

$$\mathrm{D}[af(t)+bg(t)]=af'(t)+bg'(t)=a\mathrm{D}[f(t)]+bd\mathrm{D}[g(t)].$$

Similarly, if $\mathcal{I}$ represents anti-differentiation, i.e., $\mathcal{I}[f(t)]=\int f(t)\,\mathrm{d}t$, then we again know from Calculus I that

$$\mathcal{I}[af(t)+bg(t)]=a\int f(t)\,\mathrm{d}t+b\int g(t)\,\mathrm{d}t=a\mathcal{I}[f(t)]+b\mathcal{I}[g(t)].$$

While we will not explore this issue too deeply in this text (although the idea will be used in Chapter 3.4 when discussing the solution structure for linear systems of ODEs), the implication of this fact is that much of what we study about the actions of matrices on the set of vectors also applies to operations such as differentiation and integration on the set of functions.

▷ *Remark 1.2.6.* For a simple example of a *nonlinear operator*, i.e., an operator which is not linear, consider $\mathcal{F}(x)=x^2$. We have

$$\mathcal{F}(ax+by)=(ax+by)^2=a^2x^2+2abxy+b^2y^2,$$

while

$$a\mathcal{F}(x)+b\mathcal{F}(y)=ax^2+by^2.$$

These two quantities are clearly equal for all x and y if and only if $a=b=0$; consequently, the operator $\mathcal{F}$ cannot be a linear operator.

Exercises

Exercise 1.2.1. For each of the below problems compute the product Ax when it is well-defined. If the product cannot be computed, explain why.

(a) $A=\begin{pmatrix}1&-3\\-3&2\end{pmatrix},\quad x=\begin{pmatrix}-4\\2\end{pmatrix}$

(b) $A=\begin{pmatrix}1&-2&5\\2&0&-3\end{pmatrix},\quad x=\begin{pmatrix}2\\-1\\7\end{pmatrix}$

(c) $A = \begin{pmatrix} 1 & -2 \\ 5 & 2 \\ 0 & -3 \end{pmatrix}, \quad x = \begin{pmatrix} 2 \\ -1 \\ 7 \end{pmatrix}$

(d) $A = \begin{pmatrix} 2 & -1 & -3 \end{pmatrix}, \quad x = \begin{pmatrix} 1 \\ 6 \\ -4 \end{pmatrix}.$

Exercise 1.2.2. Let

$$a_1 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 1 \\ 5 \\ 3 \end{pmatrix}, \quad b = \begin{pmatrix} -3 \\ 1 \\ 5 \end{pmatrix}.$$

Is b a linear combination of a_1, a_2, a_3? If so, are the weights unique?

Exercise 1.2.3. Let

$$A = \begin{pmatrix} 2 & 5 \\ -3 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

Is the linear system $Ax = b$ consistent? If so, what particular linear combination(s) of the columns of A give the vector b?

Exercise 1.2.4. Find all of the solutions to the homogeneous problem $Ax = 0$ when:

(a) $A = \begin{pmatrix} 1 & -3 & 6 \\ 2 & 0 & 7 \end{pmatrix}$

(b) $A = \begin{pmatrix} 1 & -3 & -4 \\ -2 & 4 & -12 \\ 0 & 2 & -4 \end{pmatrix}$

(c) $A = \begin{pmatrix} 2 & 3 & 6 \\ -3 & 5 & -1 \\ 1 & -1 & 1 \end{pmatrix}$

Exercise 1.2.5. Let

$$A = \begin{pmatrix} 2 & -1 \\ -6 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Describe the set of all vectors b for which $Ax = b$ is consistent.

Exercise 1.2.6. Determine if each of the following statements is true or false. Provide an explanation for your answer.

(a) The homogeneous system $Ax = 0$ is consistent.
(b) If b is a linear combination of a_1, a_2, then there exist unique scalars x_1, x_2 such that $b = x_1 a_1 + x_2 a_2$.
(c) If $Ax = b$ is consistent, then b is a linear combination of the rows of A.
(d) A linear combination of five vectors in $\mathbb{R}^3$ produces a vector in $\mathbb{R}^5$.
(e) In order to compute Ax, the vector x must have the same number of entries as the number of rows in A.

1.3 Sets of linear combinations of vectors

Consider the linear system

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_k).$$

This linear system is equivalent to

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k = \boldsymbol{b}.$$

The linear system is consistent if and only if the vector $\boldsymbol{b}$ is some linear combination of the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$. We now study the set of *all* linear combinations of these vectors. Once this set has been properly described, we will consider the problem of determining which of the original set of vectors are needed in order to adequately describe it.

1.3.1 Span of a set of vectors

A particular linear combination of the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ is given by $x_1\boldsymbol{a}_1 + \cdots + x_k\boldsymbol{a}_k$. The collection of all possible linear combinations of these vectors is known as the *span* of the vectors.

Span

Definition 1.3.1. Let $S = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\}$ be a set of n-vectors. The span of S,

$$\mathrm{Span}(S) = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\},$$

is the collection of all linear combinations. In other words, $\boldsymbol{b} \in \mathrm{Span}(S)$ if and only if for some $\boldsymbol{x} \in \mathbb{R}^k$,

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k.$$

The span of a collection of vectors has geometric meaning. First suppose that $\boldsymbol{a}_1 \in \mathbb{R}^3$. Recall that lines in $\mathbb{R}^3$ are defined parametrically by

$$\boldsymbol{r}(t) = \boldsymbol{r}_0 + t\boldsymbol{v},$$

where $\boldsymbol{v}$ is a vector parallel to the line and $\boldsymbol{r}_0$ corresponds to a point on the line. Since

$$\mathrm{Span}\{\boldsymbol{a}_1\} = \{t\boldsymbol{a}_1 : t \in \mathbb{R}\},$$

this set is the line through the origin which is parallel to $\boldsymbol{a}_1$.

Now suppose that $\boldsymbol{a}_1, \boldsymbol{a}_2 \in \mathbb{R}^3$ are not parallel, i.e., $\boldsymbol{a}_2 \neq c\boldsymbol{a}_1$ for some $c \in \mathbb{R}$. Set $\boldsymbol{v} = \boldsymbol{a}_1 \times \boldsymbol{a}_2$, i.e., $\boldsymbol{v}$ is a 3-vector which is perpendicular to both $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. The linearity of the dot product, and the fact that $\boldsymbol{v} \cdot \boldsymbol{a}_1 = \boldsymbol{v} \cdot \boldsymbol{a}_2 = 0$, yields

$$\boldsymbol{v} \cdot (x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2) = x_1\boldsymbol{v} \cdot \boldsymbol{a}_1 + x_2\boldsymbol{v} \cdot \boldsymbol{a}_2 = 0.$$

Thus,

$$\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\} = \{x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 : x_1, x_2 \in \mathbb{R}\}$$

is the collection of all vectors which are perpendicular to $\boldsymbol{v}$. In other words, $\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ is the plane through the origin which is perpendicular to $\boldsymbol{v}$. There are higher dimensional analogues, but unfortunately they are difficult to visualize.

Now let us consider the computation that must be done in order to determine if $\boldsymbol{b} \in \mathrm{Span}(S)$. By definition $\boldsymbol{b} \in \mathrm{Span}(S)$, i.e., $\boldsymbol{b}$ is a linear combination of the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$, if and only if there exist constants $x_1, x_2, \ldots, x_k$ such that

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k = \boldsymbol{b}.$$

Upon setting

$$\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_k), \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix},$$

by using the Definition 1.2.2 of matrix/vector multiplication we have that this condition is equivalent to solving the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. This yields:

Lemma 1.3.2. *Suppose that $S = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\}$, and set $\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_k)$. The vector $\boldsymbol{b} \in \mathrm{Span}(S)$ if and only if the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent.*

◁ *Example 1.3.3.* Letting

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} -1 \\ 2 \end{pmatrix},$$

let us determine if $\boldsymbol{b} \in \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$. As we have seen in Lemma 1.3.2, this question is equivalent to determining if the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent. Since after Gaussian elimination

$$(\boldsymbol{A}|\boldsymbol{b}) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & -4 \end{array}\right),$$

the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is equivalent to

$$x_1 = 3, \quad x_2 = -4,$$

which is easily solved. Thus, not only is $\boldsymbol{b} \in \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$, but it is the case that $\boldsymbol{b} = 3\boldsymbol{a}_1 - 4\boldsymbol{a}_2$.

◁ *Example 1.3.4.* Letting

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 2 \\ -4 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 3 \\ -1 \\ 5 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 7 \\ -7 \\ r \end{pmatrix},$$

let us determine those value(s) of r for which $\boldsymbol{b} \in \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$. As we have seen in Lemma 1.3.2, this question is equivalent to determining if the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent. Since after Gaussian elimination

$$(\boldsymbol{A}|\boldsymbol{b}) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & r-23 \end{array}\right),$$

the linear system is consistent if and only if $r = 23$. In this case $x_1 = -2, x_2 = 3$, so that $\boldsymbol{b} \in \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ with $\boldsymbol{b} = -2\boldsymbol{a}_1 + 3\boldsymbol{a}_2$.

1.3.2 Linear independence of a set of vectors

We now consider the question of how many of the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$ are needed to completely describe $\mathrm{Span}(\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_l\})$. For example, let $S = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$, where

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \ \boldsymbol{a}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \ \boldsymbol{a}_3 = \begin{pmatrix} 5 \\ -2 \\ 3 \end{pmatrix}.$$

and consider $\mathrm{Span}(S)$. If $\boldsymbol{b} \in \mathrm{Span}(S)$, then upon using Definition 1.3.1 we know there exist constants x_1, x_2, x_3 such that

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3.$$

Now, it can be checked that

$$\boldsymbol{a}_3 = 2\boldsymbol{a}_1 + 3\boldsymbol{a}_2 \quad \Leftrightarrow \quad \boldsymbol{a}_1 + 3\boldsymbol{a}_2 - \boldsymbol{a}_3 = \boldsymbol{0}, \tag{1.3.1}$$

so the vector $\boldsymbol{a}_3$ is a linear combination of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. The linear combination can be then rewritten as

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3(2\boldsymbol{a}_1 + 3\boldsymbol{a}_2) = (x_1 + 2x_3)\boldsymbol{a}_1 + (x_2 + 3x_3)\boldsymbol{a}_2.$$

In other words, the vector $\boldsymbol{b}$ is a linear combination of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ alone. Thus, the addition of $\boldsymbol{a}_3$ in the definition of $\mathrm{Span}(S)$ is superfluous, so we can write

$$\mathrm{Span}(S) = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}.$$

Since $\boldsymbol{a}_2 \neq c\boldsymbol{a}_1$ for some $c \in \mathbb{R}$, we cannot reduce the collection of vectors comprising the spanning set any further.

We say that if some nontrivial linear combination of some set of vectors produces the zero vector, such as in (1.3.1), then:

Linear dependence

Definition 1.3.5. The set of vectors $S = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\}$ is *linearly dependent* if there is a nontrivial vector $\boldsymbol{x} \neq \boldsymbol{0} \in \mathbb{R}^k$ such that

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k = \boldsymbol{0}. \tag{1.3.2}$$

Otherwise, the set of vectors is *linearly independent*.

In the preceding example the set $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ is linearly dependent, whereas the set $\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ is linearly independent.

If the set of vectors is linearly dependent, then (at least) one vector in the collection can be written as a linear combination of the other vectors (again see (1.3.1)). In particular, two vectors will be linearly dependent if and only if one is a multiple of the other. An examination of (1.3.2) through the lens of matrix/vector multiplication reveals the left-hand side is $\boldsymbol{Ax}$. Consequently, we determine if a set of vectors is linearly dependent or independent by solving the homogeneous linear system

$$\boldsymbol{Ax} = \boldsymbol{0}, \quad \boldsymbol{A} = (\boldsymbol{a}_1 \, \boldsymbol{a}_2 \cdots \boldsymbol{a}_k).$$

If there is a nontrivial solution, i.e., a solution other than the zero vector, then the vectors will be linearly dependent; otherwise, they will be independent.

Lemma 1.3.6. *Let $S = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$ be a set of n-vectors, and set $\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_k) \in \mathbb{R}^{n\times k}$. The vectors are linearly dependent if and only the linear system $\boldsymbol{Ax} = \boldsymbol{0}$ has a nontrivial solution. Alternatively, the vectors are linearly independent if and only if the only solution to $\boldsymbol{Ax} = \boldsymbol{0}$ is $\boldsymbol{x} = \boldsymbol{0}$.*

Regarding the homogeneous problem, note that if

$$\boldsymbol{Ax} = \boldsymbol{0},$$

then by the linearity of matrix/vector multiplication,

$$\boldsymbol{0} = c\boldsymbol{Ax} = \boldsymbol{A}(c\boldsymbol{x}), \quad c \in \mathbb{R}.$$

In other words, if $\boldsymbol{x}$ is a solution to the homogeneous problem, then so is $c\boldsymbol{x}$ for any constant c. Thus, if the homogeneous system has one nontrivial solution, there will necessarily be an infinite number of such solutions.

◁ *Example* 1.3.7. Let

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix}, \quad \boldsymbol{a}_3 = \begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix}, \quad \boldsymbol{a}_4 = \begin{pmatrix} -3 \\ 3 \\ -2 \end{pmatrix},$$

and consider the sets

$$S_1 = \{\boldsymbol{a}_1, \boldsymbol{a}_2\}, \quad S_2 = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}, \quad S_3 = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4\}.$$

Forming the augmented matrix and performing Gaussian elimination gives the RREF of each given matrix to be

$$(\boldsymbol{A}_1|\boldsymbol{0}) = (\boldsymbol{a}_1\ \boldsymbol{a}_2\,|\,\boldsymbol{0}) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right), \quad (\boldsymbol{A}_2|\boldsymbol{0}) = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \boldsymbol{a}_3\,|\,\boldsymbol{0}) \xrightarrow{\text{RREF}} \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

and

$$(\boldsymbol{A}_3|\boldsymbol{0}) = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \boldsymbol{a}_3\ \boldsymbol{a}_4\,|\,\boldsymbol{0}) \xrightarrow{\text{RREF}} \left(\begin{array}{cccc|c} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right).$$

Let $\mathrm{Null}(\boldsymbol{A})$ denote the set of all solutions to $\boldsymbol{Ax} = \boldsymbol{0}$ (we formalize this definition in Definition 1.4.1). We have

$$\mathrm{Null}(\boldsymbol{A}_1) = \{\boldsymbol{0}\}, \quad \mathrm{Null}(\boldsymbol{A}_2) = \mathrm{Span}\{\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}\}, \quad \mathrm{Null}(\boldsymbol{A}_3) = \mathrm{Span}\{\begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix}\}.$$

Thus, the only linearly independent set is S_1, and for the sets S_2 and S_3 we have the respective relations

$$-2\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3 = \boldsymbol{0}, \quad -2\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3 + 0\boldsymbol{a}_4 = \boldsymbol{0}.$$

In both relationships we see that $\boldsymbol{a}_3$ is a linear combination of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$,

$$\boldsymbol{a}_3 = 2\boldsymbol{a}_1 - \boldsymbol{a}_2.$$

Thus, we have

$$\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\} = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}, \quad \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4\} = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_4\}.$$

The spanning set for each of S_1 and S_2 is a plane, while the spanning set for S_3 is $\mathbb{R}^3$.

Spanning set

Definition 1.3.8. Let $S = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$, where each vector $\boldsymbol{a}_j \in \mathbb{R}^n$. We say that S is a *spanning set* for $\mathbb{R}^n$ if each $\boldsymbol{b} \in \mathbb{R}^n$ is realized as a linear combination of the vectors in S,

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k.$$

In other words, S is a spanning set if the linear system,

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_k),$$

is consistent for any $\boldsymbol{b}$.

When solving the homogeneous system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ by Gaussian elimination, it is enough to row reduce the matrix $\boldsymbol{A}$. In this case the augmented matrix $(\boldsymbol{A}|\boldsymbol{0})$ yields no additional information upon appending $\boldsymbol{0}$ to $\boldsymbol{A}$.

◁ *Example 1.3.9.* Suppose that $S = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5\}$, where each $\boldsymbol{a}_j \in \mathbb{R}^4$. Further suppose that the RREF of $\boldsymbol{A}$ is

$$\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \boldsymbol{a}_3\ \boldsymbol{a}_4\ \boldsymbol{a}_5) \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 2 & 0 & -3 \\ 0 & 1 & -1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Again letting $\mathrm{Null}(\boldsymbol{A})$ denote the set of all solutions to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, we have

$$\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}\},$$

so that

$$-2\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3 = \boldsymbol{0}, \quad -\boldsymbol{a}_2 + \boldsymbol{a}_4 = \boldsymbol{0}, \quad 3\boldsymbol{a}_1 - 2\boldsymbol{a}_2 + \boldsymbol{a}_5 = \boldsymbol{0}.$$

We have the last three vectors are each a linear combination of the first two,

$$\boldsymbol{a}_3 = 2\boldsymbol{a}_1 - \boldsymbol{a}_2, \quad \boldsymbol{a}_4 = \boldsymbol{a}_2, \quad \boldsymbol{a}_5 = -3\boldsymbol{a}_1 + 2\boldsymbol{a}_2.$$

Thus, in this case

$$\mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3,\boldsymbol{a}_4,\boldsymbol{a}_5\} = \mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2\}.$$

As we saw in these examples, the set of linearly independent vectors which form a spanning set can be found by removing from the original set those vectors which correspond to free variables in the RREF of $\boldsymbol{A}$. In the last example the free variables are x_3, x_4, x_5, which means that each of the vectors $\boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5$ is a linear combination of the first two vectors, and consequently provide no new information about the spanning set. In general, it is a relatively straightforward exercise (see Exercise 1.3.5) to show that each column of a matrix $\boldsymbol{A}$ which is not a pivot column can be written as a linear combination of the pivot columns.

Lemma 1.3.10. *Let $S = \{\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_k\}$ be a set of vectors. We have*

$$\mathrm{Span}(S) = \mathrm{Span}\{\boldsymbol{a}_{k_1},\boldsymbol{a}_{k_2},\dots,\boldsymbol{a}_{k_\ell}\},$$

where

$$\{\boldsymbol{a}_{k_1},\boldsymbol{a}_{k_2},\dots,\boldsymbol{a}_{k_\ell}\} \subset \{\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_k\}, \quad \ell \le k,$$

are those vectors which correspond to the pivot columns for the matrix $\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_k)$. Moreover, the pivot columns are a linearly independent set of vectors.

◁ *Example 1.3.11.* Let

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix}, \quad \boldsymbol{a}_3 = \begin{pmatrix} 0 \\ -7 \\ -1 \end{pmatrix},$$

and set $S = \{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$. It can be checked that

$$\boldsymbol{A} = (\boldsymbol{a}_1\,\boldsymbol{a}_2\,\boldsymbol{a}_3) \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

The first and second columns of $\boldsymbol{A}$ are the pivot columns, so

$$\mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\} = \mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2\} = \mathrm{Span}\{\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix}\}.$$

Since

$$\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix}\},$$

we have for the remaining vector,

$$-3\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3 = \boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{a}_3 = 3\boldsymbol{a}_1 - \boldsymbol{a}_2.$$

◁ *Example 1.3.12.* Suppose that $S = \{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3,\boldsymbol{a}_4\}$, where each $\boldsymbol{a}_j \in \mathbb{R}^5$. Further suppose that the RREF of $\boldsymbol{A}$ is

$$A = (a_1\, a_2\, a_3\, a_4) \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The first and third columns of A are the pivot columns, so

$$\text{Span}\{a_1, a_2, a_3, a_4\} = \text{Span}\{a_1, a_3\}.$$

Since

$$\text{Null}(A) = \text{Span}\{\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ -4 \\ 1 \end{pmatrix}\},$$

for the other two vectors we have the relationships,

$$a_2 = a_1, \quad a_4 = 3a_1 + 4a_3.$$

Exercises

Exercise 1.3.1. Determine if $b \in \text{Span}\{a_1, \dots, a_\ell\}$ for the following vectors. If the answer is YES, give the linear combination(s) which makes it true.

(a) $b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad a_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 3 \\ -5 \end{pmatrix}$

(b) $b = \begin{pmatrix} -2 \\ 5 \end{pmatrix}, \quad a_1 = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

(c) $b = \begin{pmatrix} -5 \\ -4 \\ 15 \end{pmatrix}, \quad a_1 = \begin{pmatrix} 1 \\ -1 \\ 6 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 4 \\ -1 \\ -9 \end{pmatrix}$

(d) $b = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}, \quad a_1 = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 3 \\ -1 \\ 5 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$

Exercise 1.3.2. Find the equation of the line in $\mathbb{R}^2$ which corresponds to $\text{Span}\{v_1\}$, where

$$v_1 = \begin{pmatrix} 2 \\ -5 \end{pmatrix}.$$

Exercise 1.3.3. Find the equation of the plane in $\mathbb{R}^3$ which corresponds to $\text{Span}\{v_1, v_2\}$, where

$$v_1 = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}.$$

Exercise 1.3.4. Determine if each of the following statements is true or false. Provide an explanation for your answer.

(a) The span of any two nonzero vectors in $\mathbb{R}^3$ can be viewed as a plane through the origin in $\mathbb{R}^3$.

(b) If $Ax = b$ is consistent, then $b \in \text{Span}\{a_1, a_2, \dots, a_n\}$ for $A = (a_1\, a_2 \cdots a_n)$.

(c) The number of free variables in the RREF of $\boldsymbol{A}$ is the same as the number of pivot columns.
(d) The span of a single nonzero vector in $\mathbb{R}^2$ can be viewed as a line through the origin in $\mathbb{R}^2$.

Exercise 1.3.5. Let $S = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$ be a set of vectors, and set $\boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_k)$.

(a) Show that each column of $\boldsymbol{A}$ which is not a pivot column can be written as a linear combination of the pivot columns (*Hint*: consider $\mathrm{Null}(\boldsymbol{A})$).
(b) Prove Lemma 1.3.10.

Exercise 1.3.6. Determine if the set of vectors is linearly independent. If the answer is NO, give the weights for the linear combination which results in the zero vector.

(a) $\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ -4 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} -3 \\ 12 \end{pmatrix}$
(b) $\boldsymbol{a}_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} -1 \\ 5 \end{pmatrix}$
(c) $\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{a}_3 = \begin{pmatrix} 9 \\ 6 \\ 0 \end{pmatrix}$
(d) $\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} -3 \\ -5 \\ 6 \end{pmatrix}, \quad \boldsymbol{a}_3 = \begin{pmatrix} 0 \\ 5 \\ -6 \end{pmatrix}$
(e) $\boldsymbol{a}_1 = \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}, \quad \boldsymbol{a}_3 = \begin{pmatrix} 0 \\ -11 \\ 8 \end{pmatrix}$

1.4 The structure of the solution

We now show that we can break up the solution to the consistent linear system,

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{1.4.1}$$

into two distinct pieces.

1.4.1 The homogeneous solution and the null space

As we have already seen in our discussion of linear dependence of vectors, an interesting class of linear systems which are important to solve arises when $\boldsymbol{b} = \boldsymbol{0}$:

$\mathrm{Null}(\boldsymbol{A})$ *is a nonempty set, as* $\boldsymbol{A} \cdot \boldsymbol{0} = \boldsymbol{0}$ *implies* $\{\boldsymbol{0}\} \subset \mathrm{Null}(\boldsymbol{A})$.

Null space

Definition 1.4.1. A *homogeneous linear system* is given by $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$. A *homogeneous solution*, $\boldsymbol{x}_{\mathrm{h}}$, is a solution to a homogeneous linear system. The *null space* of $\boldsymbol{A}$, denoted by $\mathrm{Null}(\boldsymbol{A})$, is the set of all solutions to a homogeneous linear system, i.e.,

$$\mathrm{Null}(\boldsymbol{A}) := \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\}.$$

Homogeneous linear systems have the important property that linear combinations of solutions are solutions; namely:

Lemma 1.4.2. *Suppose that* $\boldsymbol{x}_1,\boldsymbol{x}_2\in\mathrm{Null}(\boldsymbol{A})$, *i.e., they are two solutions to the homogeneous linear system* $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$. *It is then true that* $\boldsymbol{x}=c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2\in\mathrm{Null}(\boldsymbol{A})$ *for any* $c_1,c_2\in\mathbb{R}$; *in other words,* $\mathrm{Span}\{\boldsymbol{x}_1,\boldsymbol{x}_2\}\subset\mathrm{Null}(\boldsymbol{A})$.

Proof. The result follows immediately from the linearity of matrix/vector multiplication (see Lemma 1.2.4). In particular, we have that

$$\boldsymbol{A}(c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2)=c_1\boldsymbol{A}\boldsymbol{x}_1+c_2\boldsymbol{A}\boldsymbol{x}_2=c_1\boldsymbol{0}+c_2\boldsymbol{0}=\boldsymbol{0}. \qquad \square$$

As a consequence of the fact that linear combinations of vectors in the null space are in the null space, the homogeneous solution can be written as a linear combination of vectors, each of which resides in the null space. For example, if for a certain matrix $\boldsymbol{A}$,

$$\mathrm{Null}(\boldsymbol{A})=\mathrm{Span}\{\begin{pmatrix}-4\\2\\2\end{pmatrix},\begin{pmatrix}-1\\4\\0\end{pmatrix}\},$$

then the homogeneous solution can be written

$$\boldsymbol{x}_{\mathrm{h}}=c_1\begin{pmatrix}-4\\2\\2\end{pmatrix}+c_2\begin{pmatrix}-1\\4\\0\end{pmatrix}.$$

◁ *Example 1.4.3.* Suppose that

$$\boldsymbol{A}=\begin{pmatrix}2&-3\\-4&6\end{pmatrix}.$$

It is straightforward to check that

$$\boldsymbol{A}\begin{pmatrix}3\\2\end{pmatrix}=3\begin{pmatrix}2\\-4\end{pmatrix}+2\begin{pmatrix}-3\\6\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix},$$

so that

$$\boldsymbol{x}_1=\begin{pmatrix}3\\2\end{pmatrix}\in\mathrm{Null}(\boldsymbol{A}).$$

By Lemma 1.4.2 it is then the case that $c_1\boldsymbol{x}_1\in\mathrm{Null}(\boldsymbol{A})$ for any $c_1\in\mathbb{R}$. This can easily be checked by noting that

$$c_1\boldsymbol{x}_1 = \begin{pmatrix} 3c_1 \\ 2c_1 \end{pmatrix} \quad \Rightarrow \quad \boldsymbol{A}(c_1\boldsymbol{x}_1) = 3c_1\begin{pmatrix} 2 \\ -4 \end{pmatrix} + 2c_1\begin{pmatrix} -3 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The homogeneous solution to the homogeneous system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ is

$$\boldsymbol{x}_{\mathrm{h}} = c\begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

◁ *Example* 1.4.4. Consider the homogeneous system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, where

$$\boldsymbol{A} = \begin{pmatrix} 1 & -1 & 1 & 0 \\ -2 & 1 & -5 & -1 \\ 3 & -3 & 3 & 0 \end{pmatrix}.$$

Recall that in order to solve the linear system it is enough to put $\boldsymbol{A}$ into RREF. Using Gaussian elimination yields

$$\boldsymbol{A} \stackrel{\text{RREF}}{\longrightarrow} \begin{pmatrix} 1 & 0 & 4 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

which yields the linear system

$$x_1 + 4x_3 + x_4 = 0, \quad x_2 + 3x_3 + x_4 = 0.$$

Upon setting $x_3 = s$, $x_4 = t$ the homogeneous solution is

$$\boldsymbol{x}_{\mathrm{h}} = \begin{pmatrix} -4s - t \\ -3s - t \\ s \\ t \end{pmatrix} = s\begin{pmatrix} -4 \\ -3 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix},$$

so

$$\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} -4 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix}\}.$$

◁ *Example* 1.4.5. Consider the homogeneous system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, where

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & -3 \\ 2 & 1 & 3 & 0 \\ 1 & -1 & 0 & 3 \\ -3 & 2 & -1 & -7 \end{pmatrix}.$$

Using Gaussian elimination gives

$$\boldsymbol{A} \stackrel{\text{RREF}}{\longrightarrow} \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

which yields the linear system

$$x_1 + x_3 + x_4 = 0, \quad x_2 + x_3 - 2x_4 = 0.$$

Again setting $x_3 = s$, $x_4 = t$ the homogeneous solution is

$$\boldsymbol{x}_\text{h} = \begin{pmatrix} -s-t \\ -s+2t \\ s \\ t \end{pmatrix} = s\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix},$$

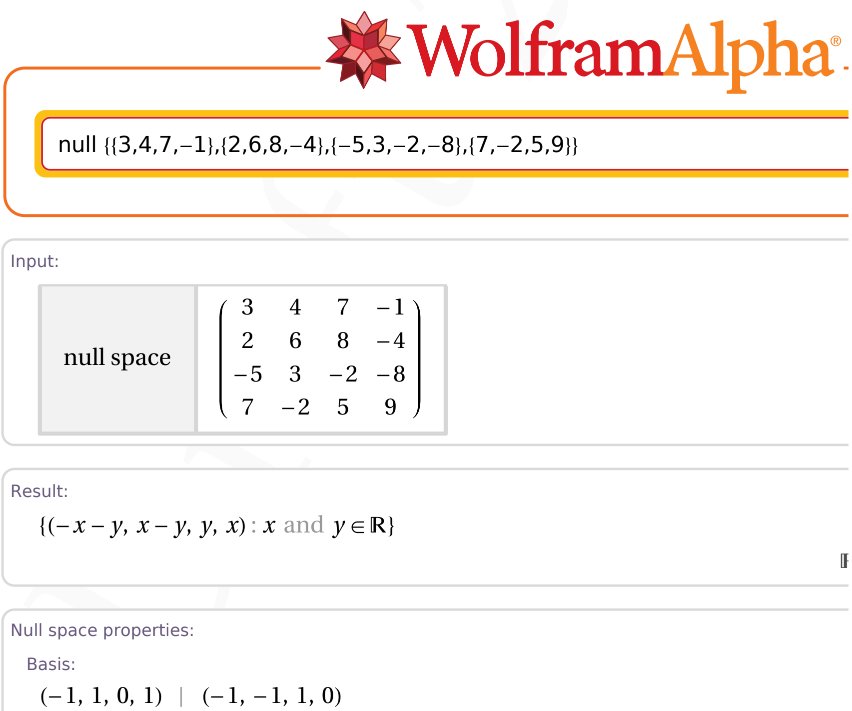
so

$$\text{Null}(\boldsymbol{A}) = \text{Span}\{\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}\}.$$

◁ *Example* 1.4.6. Consider the homogeneous linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, where

$$\boldsymbol{A} = \begin{pmatrix} 3 & 4 & 7 & -1 \\ 2 & 6 & 8 & -4 \\ -5 & 3 & -2 & -8 \\ 7 & -2 & 5 & 9 \end{pmatrix}.$$

We will use WolframAlpha to find a spanning set for Null($\boldsymbol{A}$). We have



so that

$$\text{Null}(\boldsymbol{A}) = \text{Span}\{\begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}\}.$$

1.4.2 The particular solution

Again consider the homogeneous equation $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$. We saw in Definition 1.4.1 that the homogeneous solution, $\boldsymbol{x}_{\mathrm{h}}$, resides in the null space of $\boldsymbol{A}$. Let a solution to the nonhomogeneous problem be designated as $\boldsymbol{x}_{\mathrm{p}}$ (the *particular solution*). As a consequence of the linearity of matrix/vector multiplication we have

$$\boldsymbol{A}(\boldsymbol{x}_{\mathrm{h}} + \boldsymbol{x}_{\mathrm{p}}) = \boldsymbol{A}\boldsymbol{x}_{\mathrm{h}} + \boldsymbol{A}\boldsymbol{x}_{\mathrm{p}} = \boldsymbol{0} + \boldsymbol{b} = \boldsymbol{b}.$$

In other words, the sum of the homogeneous and particular solutions, $\boldsymbol{x}_{\mathrm{h}}+\boldsymbol{x}_{\mathrm{p}}$, is a solution to the linear system (1.4.1). Indeed, any solution can be written in such a manner, simply by writing a solution $\boldsymbol{x}$ as $\boldsymbol{x} = \boldsymbol{x}_{\mathrm{h}} + (\boldsymbol{x} - \boldsymbol{x}_{\mathrm{h}})$ and designating $\boldsymbol{x}_{\mathrm{p}} = \boldsymbol{x} - \boldsymbol{x}_{\mathrm{h}}$.

Theorem 1.4.7. *All solutions to the linear system (1.4.1) are of the form*

$$\boldsymbol{x} = \boldsymbol{x}_{\mathrm{h}} + \boldsymbol{x}_{\mathrm{p}},$$

where the homogeneous solution $\boldsymbol{x}_{\mathrm{h}} \in \mathrm{Null}(\boldsymbol{A})$ is independent of $\boldsymbol{b}$, and the particular solution $\boldsymbol{x}_{\mathrm{p}}$ depends upon $\boldsymbol{b}$.

The result of Theorem 1.4.7 will be the foundation of solving not only linear systems, but also linear ordinary differential equations. It should be noted that there is a bit of ambiguity associated with the homogeneous solution. As we saw in Lemma 1.4.2, if $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathrm{Null}(\boldsymbol{A})$, then it will be the case that there is a family of homogeneous solutions given by the linear combination of these solutions, i.e., $\boldsymbol{x}_{\mathrm{h}} = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2$ for any constants $c_1, c_2 \in \mathbb{R}$. On the other hand, there really is no such ambiguity for the particular solution. Indeed, since

$$\boldsymbol{A}(c\boldsymbol{x}_{\mathrm{p}}) = c\boldsymbol{A}\boldsymbol{x}_{\mathrm{p}} = c\boldsymbol{b},$$

we have that $c\boldsymbol{x}_{\mathrm{p}}$ is a particular solution if and only if $c = 1$.

◁ *Example 1.4.8.* Consider a linear system for which

$$\boldsymbol{A} = \begin{pmatrix} 1 & 3 & 4 & -1 \\ -1 & 4 & 3 & -6 \\ 2 & -6 & -4 & 10 \\ 0 & 5 & 5 & -5 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} -2 \\ -5 \\ 8 \\ -5 \end{pmatrix}.$$

Upon performing Gaussian elimination the RREF of the augmented matrix is given by

$$(\boldsymbol{A}|\boldsymbol{b}) \xrightarrow{\text{RREF}} \left(\begin{array}{cccc|c} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

The original linear system is then equivalent to the system

$$x_1 + x_3 + 2x_4 = 1, \quad x_2 + x_3 - x_4 = -1. \tag{1.4.2}$$

The free variables for this system are x_3, x_4, so by setting $x_3 = s$ and $x_4 = t$ we get the solution to be

$$\boldsymbol{x} = \begin{pmatrix} -s-2t+1 \\ -s+t-1 \\ s \\ t \end{pmatrix} = s\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

The claim is that for the solution written in this form,

$$\boldsymbol{x}_{\rm h} = s\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{x}_{\rm p} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

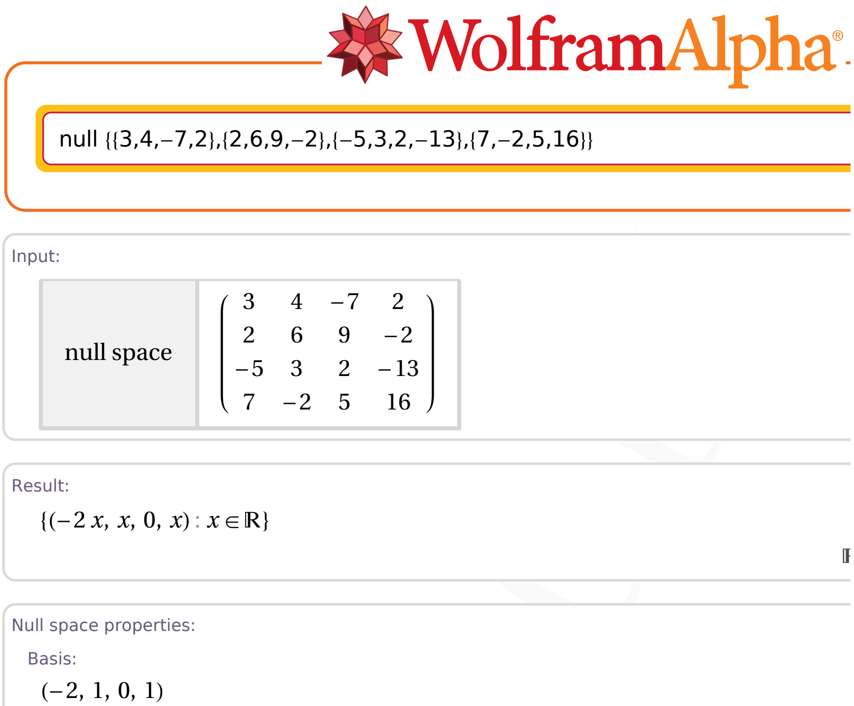
It is easy to check that $\boldsymbol{x}_{\rm p}$ is a particular solution: simply verify that $\boldsymbol{A}\boldsymbol{x}_{\rm p} = \boldsymbol{b}$. Note that $\boldsymbol{x}_{\rm p}$ is the last column of the RREF of the augmented matrix $\boldsymbol{A}|\boldsymbol{b}$). Similarly, in order to see that $\boldsymbol{x}_{\rm h}$ is a homogeneous solution, use the linearity of matrix/vector multiplication and check that

$$\boldsymbol{A}\boldsymbol{x}_{\rm h} = \boldsymbol{A}(s\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}) = s\boldsymbol{A}\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t\boldsymbol{A}\begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \boldsymbol{0}.$$

◃ *Example 1.4.9.* Consider the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, where

$$\boldsymbol{A} = \begin{pmatrix} 3 & 4 & -7 & 2 \\ 2 & 6 & 9 & -2 \\ -5 & 3 & 2 & -13 \\ 7 & -2 & 5 & 16 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 5 \\ 27 \\ 11 \\ -1 \end{pmatrix}.$$

We will use WolframAlpha to find the homogeneous solution and particular solutions. Since the homogeneous solution is a linear combination of basis vectors for the null space, we first find the null space. We have



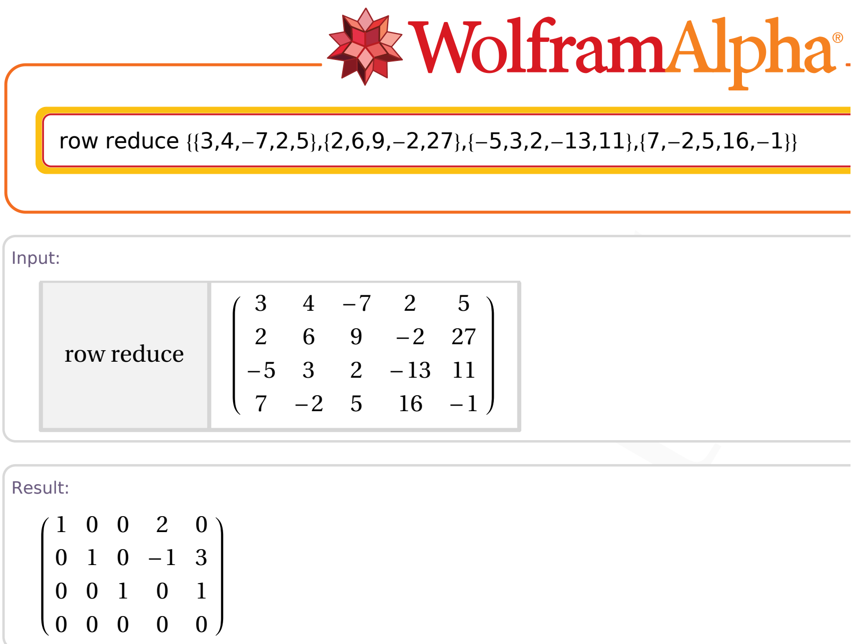
so

$$\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}\},$$

and the homogeneous solution is

$$\boldsymbol{x}_\mathrm{h} = t\begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}.$$

As for the particular solution, we compute the RREF of $(\boldsymbol{A}|\boldsymbol{b})$,



A particular solution is the last column of the RREF of $(\boldsymbol{A}|\boldsymbol{b})$, i.e.,

$$\boldsymbol{x}_{\mathrm{p}} = \begin{pmatrix} 0 \\ 3 \\ 1 \\ 0 \end{pmatrix}.$$

The solution to the full problem is the sum of the homogeneous and particular solutions,

$$\boldsymbol{x} = t\begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 1 \\ 0 \end{pmatrix}, \qquad t \in \mathbb{R}.$$

Exercises

Exercise 1.4.1. For each matrix $\boldsymbol{A}$, find $\mathrm{Null}(\boldsymbol{A})$.

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$

(b) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & -2 \\ 3 & 1 & 3 & 0 \end{pmatrix}$

(c) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 8 \\ -1 & -3 & 0 \end{pmatrix}$

(d) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 4 & 8 \end{pmatrix}$

Exercise 1.4.2. Suppose that $\boldsymbol{A} \in \mathbb{R}^{m\times n}$.

(a) Show that if $m < n$, then $\mathrm{Null}(\boldsymbol{A})$ is necessarily nontrivial; in other words, $\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell\}$ for some $\ell \geq 1$. What is a lower bound on ℓ?
(b) Give examples to show that if $m \geq n$, then $\mathrm{Null}(\boldsymbol{A})$ may or may not be trivial.

Exercise 1.4.3. For each matrix $\boldsymbol{A}$ find the general solution $\boldsymbol{x}_\mathrm{h}$ to the homogeneous problem $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$.

(a) $\begin{pmatrix} 1 & 3 \\ -2 & -6 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 3 & 8 \\ -1 & 2 & 2 \\ 3 & -4 & -2 \end{pmatrix}$

(c) $\begin{pmatrix} -2 & 1 & -5 & -6 \\ 3 & -2 & 7 & 11 \\ 4 & 5 & 17 & -16 \end{pmatrix}$

Exercise 1.4.4. For each matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$ write the solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ as $\boldsymbol{x} = \boldsymbol{x}_\mathrm{h} + \boldsymbol{x}_\mathrm{p}$, where $\boldsymbol{x}_\mathrm{h}$ is the general solution to the homogeneous problem and $\boldsymbol{x}_\mathrm{p}$ is a particular solution. Explicitly identify $\boldsymbol{x}_\mathrm{h}$ and $\boldsymbol{x}_\mathrm{p}$.

(a) $\begin{pmatrix} 1 & 3 \\ -2 & -6 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} -4 \\ 8 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 3 & 8 \\ -1 & 2 & 2 \\ 3 & -4 & -2 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 10 \\ 0 \\ 4 \end{pmatrix}$

(c) $\begin{pmatrix} -2 & 1 & -5 & -6 \\ 3 & -2 & 7 & 11 \\ 4 & 5 & 17 & -16 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 11 \\ -8 \\ 25 \\ 33 \end{pmatrix}$

Exercise 1.4.5. Given the RREF of $(\boldsymbol{A}|\boldsymbol{b})$, find the general solution. Identify the homogeneous solution, $\boldsymbol{x}_\mathrm{h}$, and particular solution, $\boldsymbol{x}_\mathrm{p}$.

(a) $\left(\begin{array}{cc|c} 1 & 0 & 7 \\ 0 & 1 & 5 \end{array}\right)$

(b) $\left(\begin{array}{ccc|c} 1 & 0 & 2 & -3 \\ 0 & 1 & -4 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$

(c) $\left(\begin{array}{ccc|c} 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 6 \end{array}\right)$

(d) $\left(\begin{array}{cccc|c} 1 & 0 & -3 & 5 & 7 \\ 0 & 1 & 1 & -2 & 9 \end{array}\right)$

(e) $\left(\begin{array}{cccc|c} 1 & 0 & 0 & -6 & 2 \\ 0 & 0 & 1 & 2 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$

1.5 Equivalence results

Before continuing we wish to summarize the results we have so far proven. Moreover, we wish to connect these results to the RREF of the appropriate matrix $\boldsymbol{A}$. We break these results into four separate pieces.

1.5.1 *A solution exists*

When we defined matrix/vector multiplication so that the linear system makes sense as $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, we showed that the linear system is consistent if and only if for some scalars $x_1, x_2, \dots, x_n \in \mathbb{R}$,

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n, \quad \boldsymbol{A} = (\boldsymbol{a}_1\,\boldsymbol{a}_2 \cdots \boldsymbol{a}_n).$$

Using Definition 1.3.1 for the span of a collection of vectors it is then the case that the system is consistent if and only if

$$\boldsymbol{b} \in \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n\}.$$

On the other hand, we solve the system by using Gaussian elimination to put the augmented matrix $(\boldsymbol{A}|\boldsymbol{b})$ into RREF. We know that the system is inconsistent if the RREF form of the augmented matrix has a row of the form $(0\,0\,0 \cdots 0|1)$; otherwise, it is consistent. These observations lead to the following equivalence result:

Theorem 1.5.1. *Regarding the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, where $\boldsymbol{A} = (\boldsymbol{a}_1\,\boldsymbol{a}_2 \cdots \boldsymbol{a}_n)$, the following are equivalent statements:*

(a) the system is consistent
(b) $\boldsymbol{b}$ is a linear combination of the columns of $\boldsymbol{A}$
(c) $\boldsymbol{b} \in \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n\}$
(d) the RREF of the augmented matrix $(\boldsymbol{A}|\boldsymbol{b})$ has no rows of the form $(0\,0\,0 \cdots 0|1)$.

◁ *Example 1.5.2.* Suppose that the coefficient matrix is given by

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}.$$

Gaussian elimination yields that the RREF of $\boldsymbol{A}$ is

$$\boldsymbol{A} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}.$$

Since the RREF of $\boldsymbol{A}$ has a zero row, the system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is not consistent for all $\boldsymbol{b}$. Since

$$\boldsymbol{A}\boldsymbol{x} = x_1\begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2\begin{pmatrix} 2 \\ 6 \end{pmatrix} = (x_1 + 2x_2)\begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

we have that

$$\mathrm{Span}\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \end{pmatrix}\} = \mathrm{Span}\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}\};$$

thus, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is by Theorem 1.5.1 consistent if and only if

$$\boldsymbol{b} \in \text{Span}\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}\} \quad \Rightarrow \quad \boldsymbol{b} = c\begin{pmatrix} 1 \\ 3 \end{pmatrix},\ c \in \mathbb{R}.$$

In other words, the vector $\boldsymbol{b}$ must be a scalar multiple of the first column of $\boldsymbol{A}$.

1.5.2 *A solution always exists*

We now wish to refine Theorem 1.5.1 in order to determine criteria which guarantee that the linear system is consistent for *any* vector $\boldsymbol{b}$. First, points (b)-(c) of Theorem 1.5.1 must be refined to say that for any $\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{b} \in \text{Span}\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_n\}$; in other words, $\text{Span}\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_n\} = \mathbb{R}^m$. Additionally, it must be the case that for a given $\boldsymbol{b}$ no row of the RREF of the augmented matrix $(\boldsymbol{A}|\boldsymbol{b})$ has row(s) of the form $(0\,0\,0 \cdots 0|0)$. Equivalently, the RREF of $\boldsymbol{A}$ must not have a zero row. If this is the case, then another vector $\boldsymbol{b}$ can be found such that the system will be inconsistent.

For example, if the RREF of the augmented system for some vector $\boldsymbol{b}_1$ is

$$(\boldsymbol{A}|\boldsymbol{b}_1) \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -3 & 5 \\ 0 & 0 & 0 \end{pmatrix},$$

then the system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}_1$ is consistent. However, for this coefficient matrix $\boldsymbol{A}$ there will exist vectors $\boldsymbol{b}_2$ such that

$$(\boldsymbol{A}|\boldsymbol{b}_2) \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

so that the system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}_2$ is not consistent.

Theorem 1.5.3. *Regarding the linear system* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, *where* $\boldsymbol{A} = (\boldsymbol{a}_1\, \boldsymbol{a}_2 \cdots \boldsymbol{a}_n)$ *and* $\boldsymbol{b} \in \mathbb{R}^m$, *the following are equivalent statements:*

(a) the system is consistent for any $\boldsymbol{b}$
(b) $\mathbb{R}^m = \text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n\}$
(c) the RREF of $\boldsymbol{A}$ *has no zero rows.*

◁ *Example* 1.5.4. Suppose that the coefficient matrix is given by

$$\boldsymbol{A} = \begin{pmatrix} 1 & -4 \\ 3 & 6 \end{pmatrix}.$$

Since Gaussian elimination yields that the RREF of $\boldsymbol{A}$ is $\boldsymbol{I}_2$, by Theorem 1.5.3 the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is consistent for any $\boldsymbol{b} \in \mathbb{R}^2$.

1.5.3 A unique solution exists

From Theorem 1.4.7 we know that all solutions are given by $\boldsymbol{x} = \boldsymbol{x}_{\mathrm{h}} + \boldsymbol{x}_{\mathrm{p}}$, where $\boldsymbol{x}_{\mathrm{h}} \in \mathrm{Null}(\boldsymbol{A})$ is a homogeneous solution and $\boldsymbol{x}_{\mathrm{p}}$ is a particular solution. Since $c\boldsymbol{x}_{\mathrm{h}} \in \mathrm{Null}(\boldsymbol{A})$ for any $c \in \mathbb{R}$ (see Lemma 1.4.2), we know that if $\boldsymbol{x}_{\mathrm{h}} \neq \boldsymbol{0}$, then the linear system has an infinite number of solutions. Since $\mathrm{Null}(\boldsymbol{A}) = \{\boldsymbol{0}\}$ if and only if the RREF of $\boldsymbol{A}$ has no free variables, a solution can be unique if and only if the RREF of $\boldsymbol{A}$ has no free variables, i.e., if every column is a pivot column. Now, following the discussion after Definition 1.3.5 we know that the columns of a matrix $\boldsymbol{A}$ are linearly independent if and only if the only solution to the homogeneous problem $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ is the trivial solution $\boldsymbol{x} = \boldsymbol{0}$. In other words, the columns are linearly independent if and only if $\mathrm{Null}(\boldsymbol{A}) = \{\boldsymbol{0}\}$. We can summarize our discussion with the following result:

Theorem 1.5.5. *The following statements about a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ are equivalent:*

(a) there is at most one solution to the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$
(b) the RREF of $\boldsymbol{A}$ has no free variables
(c) the RREF of $\boldsymbol{A}$ has a pivot position in every column
(d) the columns of $\boldsymbol{A}$ are linearly independent
(e) $\mathrm{Null}(\boldsymbol{A}) = \{\boldsymbol{0}\}$, i.e., the only solution to the homogeneous system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ is $\boldsymbol{x} = \boldsymbol{0}$.

◃ *Example 1.5.6.* Suppose that

$$\boldsymbol{A} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Since only the first two columns of $\boldsymbol{A}$ are pivot columns, and

$$\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} 5 \\ -3 \\ 1 \end{pmatrix}\} \quad \Rightarrow \quad \boldsymbol{a}_3 = -5\boldsymbol{a}_1 + 3\boldsymbol{a}_2,$$

solutions to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, if they exist, cannot be unique. Moreover, since the RREF of $\boldsymbol{A}$ has a zero row, the linear system will not always be consistent. In order for the system to be consistent, it is necessary that $\boldsymbol{b} \in \mathrm{Span}(\{\boldsymbol{a}_1, \boldsymbol{a}_2\})$.

◃ *Example 1.5.7.* Suppose that $\boldsymbol{A} \in \mathbb{R}^{6\times 4}$ has 4 pivot columns. Since that is the maximal number of pivot columns, by Theorem 1.5.5 the columns are a linearly independent set. Consequently, the solution to a consistent linear system, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, will be unique. The columns do not form a spanning set for $\mathbb{R}^6$, however, since the RREF of $\boldsymbol{A}$ will have two zero rows.

◃ *Example 1.5.8.* Suppose that $\boldsymbol{A} \in \mathbb{R}^{3\times 5}$. Since $\boldsymbol{A}$ has more columns than rows, it is impossible for the RREF of $\boldsymbol{A}$ to have a pivot position in every column. Indeed, the RREF of $\boldsymbol{A}$ must have at least two free variables, and there can be no more than three pivot columns. Hence, by Theorem 1.5.5 the columns of $\boldsymbol{A}$ cannot be linearly independent. We cannot say that the columns form a spanning set for $\mathbb{R}^3$ without knowing something more about the RREF of $\boldsymbol{A}$. If we are told that the RREF of $\boldsymbol{A}$ has two pivot positions, then it must be the case that the RREF of $\boldsymbol{A}$ has one zero row; hence, by Theorem 1.5.3 the columns cannot form a spanning set. However, if we are told that the RREF of $\boldsymbol{A}$ has three pivot positions (the maximum number possible), then it must be the case that the

RREF of $\boldsymbol{A}$ has no zero rows, which by Theorem 1.5.3 means that the columns do indeed form a spanning set. In any case, the existence of free variables means that there will be an infinite number of solutions to any consistent linear system, $\boldsymbol{Ax} = \boldsymbol{b}$.

1.5.4 A unique solution always exists

We finally consider the problem of determining when there will always be a unique solution to the linear system. By Theorem 1.5.5(c) the existence of a unique solution requires that every column be a pivot column. This is possible if and only if the RREF of $\boldsymbol{A}$ has no free variables. In order for the RREF of $\boldsymbol{A}$ to have no free variables, it must be the case that the number of rows is greater than or equal to the number of columns. On the other hand, by Theorem 1.5.3(c) the existence of a solution requires that the RREF of $\boldsymbol{A}$ have no zeros rows. The lack of zero rows in the RREF of $\boldsymbol{A}$ is possible if and only if the number of rows is greater than or equal to the number of columns. In conclusion, we see that it is possible to always have a unique solution if and only if the number of rows is equal to the number of columns, i.e., if the matrix is square.

Henceforth assume that $\boldsymbol{A}$ is square. The RREF of $\boldsymbol{A}$ can have free variables if and only if the RREF of $\boldsymbol{A}$ has zero rows. If the RREF of $\boldsymbol{A}$ has no zero rows, then since it is square it must be the case that:

(a) the RREF of $\boldsymbol{A}$ is the identity matrix $\boldsymbol{I}_n$
(b) the columns of $\boldsymbol{A}$ are linearly independent.

By Theorem 1.5.1 the lack of zero rows for the RREF of $\boldsymbol{A}$ means that the system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for any $\boldsymbol{b}$, and by Theorem 1.5.3 this lack of zero rows implies that the columns of $\boldsymbol{A}$ form a spanning set for $\mathbb{R}^n$.

Theorem 1.5.9. *The following statements about a square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ are equivalent:*

(a) there is only one solution to the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ for any $\boldsymbol{b}$
(b) the RREF of $\boldsymbol{A}$ is $\boldsymbol{I}_n$
(c) the columns of $\boldsymbol{A}$ are linearly independent
(d) the columns of $\boldsymbol{A}$ form a spanning set for $\mathbb{R}^n$
(e) $\mathrm{Null}(\boldsymbol{A}) = \{\boldsymbol{0}\}$.

◁ *Example 1.5.10.* We have

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 9 \\ -1 & 4 & 3 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Since the RREF of $\boldsymbol{A}$ is not the identity, $\boldsymbol{I}_3$, the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is not always consistent. If it is consistent, since

$$\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}\} \quad\Rightarrow\quad \boldsymbol{a}_3 = \boldsymbol{a}_1 + \boldsymbol{a}_2,$$

the columns of $\boldsymbol{A}$ are not linearly independent, so the solution will not be unique.

When discussing linear dependence we can use Definition 1.3.5 in a more general sense. Suppose that $\{f_1(t), f_2(t), \dots, f_k(t)\}$ is a set of real-valued functions, each of which has at least $k-1$ continuous derivatives. We say that these functions are linearly dependent on the interval $a < t < b$ if there is a nontrivial vector $\boldsymbol{x} \in \mathbb{R}^k$ such that

$$x_1 f_1(t) + x_2 f_2(t) + \cdots + x_k f_k(t) \equiv 0, \quad a < t < b.$$

How do we determine if this set of functions is linearly dependent? The problem is that unlike the previous examples it is not at all clear how to formulate this problem as a homogeneous linear system.

We overcome this difficulty in the following manner. Suppose that the functions are linearly dependent. Since the linear combination of the functions is identically zero, it will be the case that a derivative of the linear combination will also be identically zero, i.e.,

$$x_1 f_1'(t) + x_2 f_2'(t) + \cdots + x_k f_k'(t) \equiv 0, \quad a < t < b.$$

We can take a derivative of the above to then get

$$x_1 f_1''(t) + x_2 f_2''(t) + \cdots + x_k f_k''(t) \equiv 0, \quad a < t < b,$$

and continuing in the fashion we have for $j = 0, \dots, k-1$,

$$x_1 f_1^{(j)}(t) + x_2 f_2^{(j)}(t) + \cdots + x_k f_k^{(j)}(t) \equiv 0, \quad a < t < b.$$

We have now derived a system of k linear equations, which is given by

$$\boldsymbol{W}(t)\boldsymbol{x} \equiv \boldsymbol{0}, \quad \boldsymbol{W}(t) := \begin{pmatrix} f_1(t) & f_2(t) & \cdots & f_k(t) \\ f_1'(t) & f_2'(t) & \cdots & f_k'(t) \\ f_1''(t) & f_2''(t) & \cdots & f_k''(t) \\ \vdots & \vdots & \vdots & \vdots \\ f_1^{(k-1)}(t) & f_2^{(k-1)}(t) & \cdots & f_k^{(k-1)}(t) \end{pmatrix}.$$

The matrix $\boldsymbol{W}(t)$ is known as the *Wronskian* for the set of functions $\{f_1(t), f_2(t), \dots, f_k(t)\}$.

We now see that the functions will be linearly dependent if there is a nontrivial vector $\boldsymbol{x}$, which does *not* depend on t, such that $\boldsymbol{W}(t)\boldsymbol{x} = \boldsymbol{0}$ for each $a < t < b$. Conversely, the functions will be linearly independent if there is (at least) one $a < t_0 < b$ such that the only solution to $\boldsymbol{W}(t_0)\boldsymbol{x} = \boldsymbol{0}$ is the trivial solution $\boldsymbol{x} = \boldsymbol{0}$. In other words, upon invoking Theorem 1.5.9(e) we see that the functions will be linearly independent if there is (at least) one value of t_0 such that the RREF of $\boldsymbol{W}(t_0)$ is the identity matrix $\boldsymbol{I}_k$.

◁ *Example* 1.5.11. For a concrete example, consider the set $\{1, t, t^2, t^3\}$ on the interval $-\infty < t < +\infty$. The Wronskian associated with this set of functions is

$$\boldsymbol{W}(t) = \begin{pmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 2 & 6t \\ 0 & 0 & 0 & 6 \end{pmatrix}.$$

It is clear that

$$\boldsymbol{W}(0) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix} \xrightarrow{\text{RREF}} \boldsymbol{I}_4,$$

which by the above discussion implies that the set of functions is linearly independent.

◃ *Example 1.5.12.* For another example, consider the set $\{\sin(t), \cos(t)\}$ on the interval $0 \le t \le 2\pi$. The Wronskian for this set of functions is

$$W(t) = \begin{pmatrix} \sin(t) & \cos(t) \\ \cos(t) & -\sin(t) \end{pmatrix}.$$

It is clear that

$$\boldsymbol{W}(\pi/2) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \xrightarrow{\text{RREF}} \boldsymbol{I}_2,$$

so the set of functions is linearly independent.

Exercises

Exercise 1.5.1. Suppose that the RREF of $\boldsymbol{A} \in \mathbb{R}^{5\times5}$ has one zero row.

(a) Is $\boldsymbol{Ax} = \boldsymbol{b}$ consistent for any $\boldsymbol{b} \in \mathbb{R}^5$? Why, or why not?
(b) If $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent, how many solutions are there?

Exercise 1.5.2. Suppose that the RREF of $\boldsymbol{A} \in \mathbb{R}^{9\times9}$ has seven pivot columns.

(a) Is $\boldsymbol{Ax} = \boldsymbol{b}$ consistent for any $\boldsymbol{b} \in \mathbb{R}^9$? Why, or why not?
(b) If $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent, how many solutions are there?

Exercise 1.5.3. Determine if each of the following statements is true or false. Provide an explanation for your answer.

(a) If $\boldsymbol{A} \in \mathbb{R}^{5\times3}$, then it is possible for the columns of $\boldsymbol{A}$ to span $\mathbb{R}^3$.
(b) If the RREF of $\boldsymbol{A} \in \mathbb{R}^{9\times7}$ has three zeros rows, then $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for any vector $\boldsymbol{b} \in \mathbb{R}^9$.
(c) If $\boldsymbol{A} \in \mathbb{R}^{5\times9}$, then $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for any $\boldsymbol{b} \in \mathbb{R}^5$.
(d) If the RREF of $\boldsymbol{A} \in \mathbb{R}^{12\times16}$ has 12 pivot columns, then $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for any $\boldsymbol{b} \in \mathbb{R}^{12}$.
(e) If $\boldsymbol{Av}_j = \boldsymbol{0}$ for $j = 1, 2$, then $x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 \in \mathrm{Null}(\boldsymbol{A})$ for any $x_1, x_2 \in \mathbb{R}$.
(f) If $\boldsymbol{A} \in \mathbb{R}^{5\times7}$ is such that $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for every vector $\boldsymbol{b} \in \mathbb{R}^5$, then the RREF of $\boldsymbol{A}$ has at least one zero row.
(g) If $\boldsymbol{A} \in \mathbb{R}^{7\times6}$, then $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for any $\boldsymbol{b} \in \mathbb{R}^7$.

Exercise 1.5.4. For the given set S, determine whether the set is linearly dependent or linearly independent.

(a) $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$, where

$$\boldsymbol{v}_1 = \begin{pmatrix} 3 \\ -4 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

(b) $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$, where

$$\boldsymbol{v}_1 = \begin{pmatrix} 2 \\ -5 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} -4 \\ 10 \end{pmatrix}.$$

(c) $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$, where

$$\boldsymbol{v}_1 = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 2 \\ -6 \end{pmatrix}, \quad \boldsymbol{v}_3 = \begin{pmatrix} 9 \\ 15 \\ 7 \end{pmatrix}.$$

(d) $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$, where

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 3 \\ -2 \\ 5 \end{pmatrix}, \quad \boldsymbol{v}_3 = \begin{pmatrix} -7 \\ 1 \\ 4 \end{pmatrix}, \quad \boldsymbol{v}_4 = \begin{pmatrix} 0 \\ -5 \\ 8 \end{pmatrix}.$$

Exercise 1.5.5. Set

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ -2 \\ -5 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix}, \quad \boldsymbol{v}_3 = \begin{pmatrix} 8 \\ -7 \\ r \end{pmatrix}.$$

(a) For which value(s) of r are the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ linearly independent?
(b) For which value(s) of r is $\boldsymbol{v}_3 \in \mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$?
(c) How are (a) and (b) related?

Exercise 1.5.6. Suppose that $\boldsymbol{A} \in \mathbb{R}^{5\times 9}$, and further suppose that the RREF of $\boldsymbol{A}$ has five free variables.

(a) Do the columns of $\boldsymbol{A}$ span $\mathbb{R}^5$? Explain.
(b) Are the columns of $\boldsymbol{A}$ linearly dependent, linearly independent, or is not possible to say without more information? Explain.

Exercise 1.5.7. Suppose that $\boldsymbol{A} \in \mathbb{R}^{7\times 4}$, and further suppose that the RREF of $\boldsymbol{A}$ has zero free variables.

(a) Do the columns of $\boldsymbol{A}$ span $\mathbb{R}^7$? Explain.
(b) Are the columns of $\boldsymbol{A}$ linearly dependent, linearly independent, or is not possible to say without more information? Explain.

Exercise 1.5.8. Suppose that $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. For what relationship between m and n will it be necessarily true that:

(a) $\mathrm{Null}(\boldsymbol{A})$ is nontrivial.
(b) the columns of $\boldsymbol{A}$ do not span $\mathbb{R}^m$.

Exercise 1.5.9. Determine if each of the following statements is true or false. Provide an explanation for your answer.

(a) If $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ with $m > n$, then the columns of $\boldsymbol{A}$ must be linearly independent.
(b) If $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ has a pivot in every column, then the columns of $\boldsymbol{A}$ span $\mathbb{R}^m$.
(c) If $\mathrm{Null}(\boldsymbol{A})$ is nontrivial, then the columns of $\boldsymbol{A}$ are linearly independent.
(d) If $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ with $m \neq n$, then it is possible for the columns of $\boldsymbol{A}$ to both span $\mathbb{R}^m$ and be linearly independent.

Exercise 1.5.10. Show that the following sets of functions are linearly independent:

(a) $\{\mathrm{e}^t, \mathrm{e}^{2t}, \mathrm{e}^{3t}\}$, where $-\infty < t < +\infty$
(b) $\{1, \cos(t), \sin(t)\}$, where $-\infty < t < +\infty$
(c) $\{\mathrm{e}^t, t\mathrm{e}^t, t^2\mathrm{e}^t, t^3\mathrm{e}^t\}$, where $-\infty < t < +\infty$
(d) $\{1, t, t^2, \dots, t^k\}$ for any $k \geq 4$, where $-\infty < t < +\infty$
(e) $\{\mathrm{e}^{at}, \mathrm{e}^{bt}\}$ for $a \neq b$, where $-\infty < t < +\infty$

1.6 Subspaces

The null space of $\boldsymbol{A}$, $\mathrm{Null}(\boldsymbol{A})$, is an important example of a more general set:

If S is a subspace, then $\boldsymbol{0} \in S$.

Subspace

Definition 1.6.1. A nonempty set $S \subset \mathbb{R}^n$ is a *subspace* if

$$\boldsymbol{x}_1, \boldsymbol{x}_2 \in S \quad \Rightarrow \quad c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 \in S, \quad c_1, c_2 \in \mathbb{R}.$$

A subspace S is a subset of a *vector space*. A *real* vector space, V, is a collection of elements, call vectors, on which are defined two operations, addition and scalar multiplication by real numbers. If $\boldsymbol{x}, \boldsymbol{y} \in V$, then $c_1\boldsymbol{x} + c_2\boldsymbol{y} \in V$ for any real scalars c_1, c_2. The following axioms must also be satisfied:

(a) commutativity of vector addition: $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$
(b) associativity of vector addition: $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$
(c) existence of an additive identity: there exists a $\boldsymbol{0} \in V$ such that $\boldsymbol{x} + \boldsymbol{0} = \boldsymbol{x}$
(d) existence of an additive inverse: for each $\boldsymbol{x}$ there exists a $\boldsymbol{y}$ such that $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{0}$
(e) existence of a multiplicative identity: $1 \cdot \boldsymbol{x} = \boldsymbol{x}$
(f) first multiplicative distributive law: $c(\boldsymbol{x} + \boldsymbol{y}) = c\boldsymbol{x} + c\boldsymbol{y}$
(g) second multiplicative distributive law: $(c_1 + c_2)\boldsymbol{x} = c_1\boldsymbol{x} + c_2\boldsymbol{x}$
(h) relation to ordinary multiplication: $(c_1c_2)\boldsymbol{x} = c_1(c_2\boldsymbol{x}) = c_2(c_1\boldsymbol{x})$.

The set of all vectors with complex-valued coefficients, $\mathbb{C}^n$, is a vector space (see Chapter 1.12 if you are not familiar with complex numbers). In this case the constants c_1 and c_2 are complex-valued.

Examples of vector spaces include:

(a) the set of n-vectors, $\mathbb{R}^n$,
(b) the set of matrices $\boldsymbol{R}^{m\times n}$
(c) the set of all polynomials of degree n.

Going back to subspaces of $\mathbb{R}^n$, which is all we will (primarily) be concerned with in this text, by using the Definition 1.6.1 we see that that the span of a collection of vectors is a subspace:

Lemma 1.6.2. *The set $S = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$ is a subspace.*

Proof. Suppose that $\boldsymbol{b}_1, \boldsymbol{b}_2 \in S$. By Lemma 1.3.2 there exist vectors $\boldsymbol{x}_1, \boldsymbol{x}_2$ such that for $\boldsymbol{A} = (\boldsymbol{a}_1\, \boldsymbol{a}_2 \cdots \boldsymbol{a}_k)$,

$$\boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{b}_1, \quad \boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{b}_2.$$

We must now show that for the vector $\boldsymbol{b} = c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2$ there is a vector $\boldsymbol{x}$ such that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, as it will then be true that $\boldsymbol{b} \in S$. However, if we choose $\boldsymbol{x} = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2$, then by the linearity of matrix/vector multiplication we have that

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}(c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2) = c_1\boldsymbol{A}\boldsymbol{x}_1 + c_2\boldsymbol{A}\boldsymbol{x}_2 = c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 = \boldsymbol{b}.$$ □

While we will not prove it here, it can be shown that any subspace of $\mathbb{R}^n$ is realized as the span of some collection of vectors in $\mathbb{R}^n$. In other words, in the vector space $\mathbb{R}^n$ there are no other subspaces other than those given in Lemma 1.6.2.

Lemma 1.6.3. *If $S \subset \mathbb{R}^n$ is a subspace, then there is a collection of vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$ such that $S = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$.*

◁ *Example* 1.6.4. Suppose that

$$S = \{\begin{pmatrix} x_1 + 2x_2 \\ -3x_2 \\ 4x_1 + x_2 \end{pmatrix} : x_1, x_2 \in \mathbb{R}\}.$$

Since

$$\begin{pmatrix} x_1 + 2x_2 \\ -3x_2 \\ 4x_1 + x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} = \mathrm{Span}\{\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}\},$$

by Lemma 1.6.2 the set is a subspace.

◁ *Example* 1.6.5. Suppose that

$$S = \{\begin{pmatrix} x_1 + 2x_2 \\ 1 - 3x_2 \\ 4x_1 + x_2 \end{pmatrix} : x_1, x_2 \in \mathbb{R}\}.$$

We have that $\boldsymbol{b} \in S$ if and only if

$$\boldsymbol{b} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_1 \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}.$$

Since

$$2\boldsymbol{b} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + 2x_1 \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} + 2x_2 \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix},$$

it is clear that $2\boldsymbol{b} \notin S$. Consequently, S cannot be a subspace. Alternatively, it is easy to show that $\boldsymbol{0} \notin S$.

Another important example of a subspace which is directly associated with a matrix is the column space:

Column space

Definition 1.6.6. The *column space* of a matrix $\boldsymbol{A}$, $\mathrm{Col}(\boldsymbol{A})$, is the set of all linear combinations of the columns of $\boldsymbol{A}$.

The column space is also known as the range of $\boldsymbol{A}$, $R(\boldsymbol{A})$.

Setting $\boldsymbol{A} = (\boldsymbol{a}_1\, \boldsymbol{a}_1 \cdots \boldsymbol{a}_k)$, we can rewrite the column space as

$$\mathrm{Col}(\boldsymbol{A}) = \{x_1\boldsymbol{a}_1 + x_1\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k : x_1, x_2, \dots, x_k \in \mathbb{R}\} = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}.$$

By Lemma 1.6.2 $\mathrm{Col}(\boldsymbol{A})$ is a subspace. Furthermore, if $\boldsymbol{b} \in \mathrm{Col}(\boldsymbol{A})$, then it must be the case that for some weights $x_1, \dots, x_k$,

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k \quad \Rightarrow \quad \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}.$$

The formulation on the right follows from the definition of matrix/vector multiplication. This gives us:

Lemma 1.6.7. $\mathrm{Col}(\boldsymbol{A})$ *is a subspace, and the column space has the equivalent definition*

$$\mathrm{Col}(\boldsymbol{A}) = \{\boldsymbol{b} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \text{ is consistent}\}.$$

With these notions in mind, we can revisit the statement of Theorem 1.5.1 in order to make an equivalent statement. Theorem 1.5.1(c) states that a linear system is consistent if and only if the vector $\boldsymbol{b}$ is in the span of the column vectors of the matrix $\boldsymbol{A}$. The definition of the column space and the formulation of Lemma 1.6.7 yields the following restatement of Theorem 1.5.1:

Theorem 1.6.8. *Regarding the linear system* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$*, where* $\boldsymbol{A} = (\boldsymbol{a}_1\, \boldsymbol{a}_2 \cdots \boldsymbol{a}_n)$*, the following are equivalent statements:*

(a) the system is consistent
(b) $\boldsymbol{b} \in \mathrm{Col}(\boldsymbol{A})$
(c) the RREF of the augmented matrix $(\boldsymbol{A}|\boldsymbol{b})$ *has no rows of the form* $(0\,0\,0 \cdots 0|1)$.

Exercises

Exercise 1.6.1. Set

$$S = \{\begin{pmatrix} 2s-3t \\ -s+4t \\ 7t \end{pmatrix} : s,t \in \mathbb{R}\} \subset \mathbb{R}^3.$$

Is S a subspace? If so, determine vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k$ such that $S = \mathrm{Span}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k\}$, and determine $\dim[S]$. Otherwise, explain why S is not a subspace.

Exercise 1.6.2. Set

$$S = \{\begin{pmatrix} 4s+2t \\ 1-3s-t \\ s+9t \end{pmatrix} : s,t \in \mathbb{R}\} \subset \mathbb{R}^3.$$

Is S a subspace? If so, determine vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k$ such that $S = \mathrm{Span}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k\}$, and determine $\dim[S]$. Otherwise, explain why S is not a subspace.

Exercise 1.6.3. Let $A, B \subset \mathbb{R}^n$ be subspaces, and define

$$A + B = \{\boldsymbol{x} : \boldsymbol{x} = \boldsymbol{a} + \boldsymbol{b},\ \boldsymbol{a} \in A,\ \boldsymbol{b} \in B\}.$$

Show that $A + B$ is a subspace. (*Hint*: use the fact that set is a subspace if and only if it is the span of a collection of vectors)

Exercise 1.6.4. Show that the set of 2×2 matrices, $\mathbb{R}^{2\times 2}$, is a vector space.

Exercise 1.6.5. Consider the set of matrices in $\mathbb{R}^{2\times 2}$ given by

$$S = \{\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a + d = 0\}.$$

Show that S is a subspace. (*Hint*: find a set of matrices such that any $\boldsymbol{A} \in S$ can be written as a linear combination of these matrices)

1.7 Basis and dimension of a subspace

1.7.1 Basis

The next question to consider is the "size" of a subspace. The number that we assign to the "size" should reflect the intuition that a plane in $\mathbb{R}^3$ is bigger than a line in $\mathbb{R}^3$, and hence the number assigned to a plane should be larger than the number associated with the line. Regarding a given plane going through the origin in $\mathbb{R}^3$, while the geometric object itself is unique, there are many ways to describe it. For example, in Calculus we learned that it can be described as being the set of all vectors which are perpendicular to a certain vector. Conversely, we could describe it as the span of a collection of vectors which lie in the plane. The latter notion is the one that we will use, as it more easily generalizes to higher dimensions. One way to determine the "size" of the subspace is to then count the number of spanning vectors. Because an arbitrarily high number of vectors could be used as the spanning set, in order to uniquely determine the size of the space we must restrict the possible number of spanning vectors as much as possible. This restriction requires that we use only the *linearly independent* vectors (see Definition 1.3.5) in the spanning set. We first label these vectors:

Basis

Definition 1.7.1. A set $B = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$ is a *basis* for a subspace S if

(a) the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k$ are linearly independent
(b) $S = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$.

In other words, the set of vectors is a basis if

(a) any vector in S can be written as a linear combination of the basis vectors
(b) there are not so many vectors in the set that (at least) one of them can be written as a linear combination of the others.

◁ *Example 1.7.2.* We wish to find a basis for the column space, $\mathrm{Col}(\boldsymbol{A})$, of the matrix

$$\boldsymbol{A} = \begin{pmatrix} -2 & 3 & -1 \\ 3 & -5 & 1 \\ 6 & -7 & 5 \end{pmatrix}.$$

Since

$$\boldsymbol{A} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

we have

$$\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix}\} \quad \Rightarrow \quad \boldsymbol{a}_3 = 2\boldsymbol{a}_1 + \boldsymbol{a}_2.$$

Consequently, regarding the column space we have

$$\mathrm{Col}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} -2 \\ 3 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ -5 \\ -7 \end{pmatrix}\}.$$

By Lemma 1.3.10 these first two columns are linearly independent. In conclusion, these two pivot columns are a basis for Col($\boldsymbol{A}$).

The previous example points us towards a general truth. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ be given. By Lemma 1.3.10 the pivot columns of $\boldsymbol{A}$ form a linearly independent spanning set for Col($\boldsymbol{A}$). Thus, by the definition of basis we have:

Lemma 1.7.3. *The pivot columns of $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ form a basis for the column space,* Col($\boldsymbol{A}$).

A basis for a subspace is not unique. For example,

$$B_1 = \{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad B_2 = \{\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\0\end{pmatrix}\},$$

are each a basis for the x_1x_2-plane in $\mathbb{R}^3$. However, we do have the intuitive result that the *number* of basis vectors for a subspace is unique.

Lemma 1.7.4. *If $A = \{\boldsymbol{a}_1, \dots, \boldsymbol{a}_k\}$ and $B = \{\boldsymbol{b}_1, \dots, \boldsymbol{b}_m\}$ are two bases for a subspace S, then it must be the case that $k = m$. In other words, all bases for a subspace have the same number of vectors.*

Proof. The result is geometrically intuitive. The mathematical proof is as follows. Start by forming the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ via

$$\boldsymbol{A} = (\boldsymbol{a}_1\, \boldsymbol{a}_2 \cdots \boldsymbol{a}_k), \quad \boldsymbol{B} = (\boldsymbol{b}_1\, \boldsymbol{b}_2 \cdots \boldsymbol{b}_m).$$

The columns of each matrix are linearly independent, so by Theorem 1.5.5(e) the null space of each matrix is trivial,

$$\text{Null}(\boldsymbol{A}) = \{\boldsymbol{0}\}, \quad \text{Null}(\boldsymbol{B}) = \{\boldsymbol{0}\}.$$

Since A is a basis, each vector in B is a linear combination of the vectors in A; in particular, for each $\boldsymbol{b}_j$ there is a vector $\boldsymbol{c}_j$ such that

$$\boldsymbol{b}_j = \boldsymbol{A}\boldsymbol{c}_j, \quad j = 1, \dots, m.$$

If we set

$$\boldsymbol{C} = (\boldsymbol{c}_1\, \boldsymbol{c}_2 \cdots \boldsymbol{c}_m) \in \mathbb{R}^{k\times m},$$

the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are then related by

$$\boldsymbol{B} = (\boldsymbol{b}_1\, \boldsymbol{b}_2 \cdots \boldsymbol{b}_m) = (\boldsymbol{A}\boldsymbol{c}_1\, \boldsymbol{A}\boldsymbol{c}_2 \cdots \boldsymbol{A}\boldsymbol{c}_m) = \boldsymbol{A}\boldsymbol{C}$$

(see Chapter 1.9 for the formal definition of matrix/matrix multiplication).

Suppose that $\boldsymbol{s} \in S$. Since A and B are each a basis there exist unique vectors $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ such that

$$\boldsymbol{s} = \boldsymbol{A}\boldsymbol{x}_A = \boldsymbol{B}\boldsymbol{x}_B.$$

But, the relation $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{C}$ implies that

$$\boldsymbol{A}\boldsymbol{x}_A = (\boldsymbol{A}\boldsymbol{C})\boldsymbol{x}_B = \boldsymbol{A}(\boldsymbol{C}\boldsymbol{x}_B) \quad \Rightarrow \quad \boldsymbol{A}(\boldsymbol{x}_A - \boldsymbol{C}\boldsymbol{x}_B) = \boldsymbol{0}$$

(the last implicated equality follows from the linearity of matrix/matrix multiplication). Recalling that $\text{Null}(\boldsymbol{A}) = \{\boldsymbol{0}\}$, $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ are related via

$$\boldsymbol{A}(\boldsymbol{x}_A - \boldsymbol{C}\boldsymbol{x}_B) = \boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{x}_A - \boldsymbol{C}\boldsymbol{x}_B = \boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{x}_A = \boldsymbol{C}\boldsymbol{x}_B.$$

Since $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ are unique, this linear system must have only unique solutions, which means that $\text{Null}(\boldsymbol{C}) = \{\boldsymbol{0}\}$ (again see Theorem 1.5.5(e)). Thus, the RREF of $\boldsymbol{C}$ has no free variables, which implies that $\boldsymbol{C}$ has at least as many rows as columns, i.e., $m \le k$.

In order to get the reverse inequality, and thus achieve the desired equality, we first write $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{D}$ for some $\boldsymbol{D} \in \mathbb{R}^{m\times k}$. Following the same argument as above, we eventually conclude that $\boldsymbol{D}\boldsymbol{x}_A = \boldsymbol{x}_B$ has unique solutions only, which is possible only if $k \le m$. We conclude by noting that $m \le k$ and $k \le m$ if and only if $k = m$, which is the desired result. □

1.7.2 Dimension

Because the number of vectors in a basis of a subspace is fixed, this quantity gives a good way to describe the "size" of a subspace.

Dimension

Definition 1.7.5. If $B = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$ is a basis for a subspace S, then the *dimension of S*, $\dim[S]$, is the number of basis vectors:

$$\dim(S) = k.$$

◁ *Example* 1.7.6. Let $\boldsymbol{e}_j$ for $j = 1, \dots, n$ denote the j^{th} column vector in the identity matrix $\boldsymbol{I}_n$. Since $\boldsymbol{I}_n$ is in RREF, by Theorem 1.5.9 the set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ is linearly independent and forms a spanning set for $\mathbb{R}^n$; in other words, it is a basis for $\mathbb{R}^n$. By Definition 1.7.5 we then have the familiar result that $\dim[\mathbb{R}^n] = n$.

Regarding the column space and its dimension, we use the following moniker:

Rank

Definition 1.7.7. The dimension of the column space of a matrix is known as the *rank*,

$$\text{rank}(\boldsymbol{A}) = \dim[\text{Col}(\boldsymbol{A})].$$

We now relate the column space to the null space through their dimensions. The pivot columns of the RREF of $\boldsymbol{A}$ are a basis for $\text{Col}(\boldsymbol{A})$ (see Lemma 1.7.3), so

$$\text{rank}(\boldsymbol{A}) = \#\text{ of pivot columns}.$$

In addition, the number of free variables is the number of linearly independent vectors that form a spanning set for $\text{Null}(\boldsymbol{A})$. Consequently, we can say

$$\dim[\text{Null}(\boldsymbol{A})] = \#\text{ of free variables}.$$

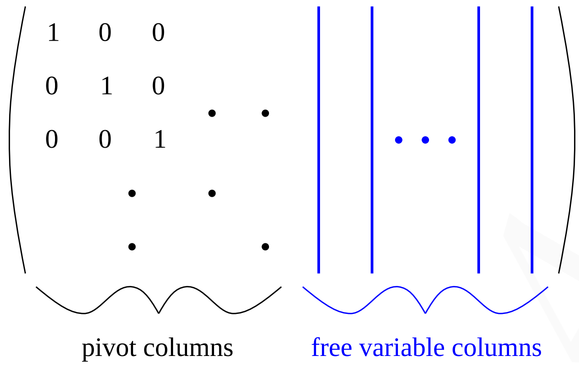


Fig. 1.2 (color online) A cartoon of a matrix in RREF. A vertical (blue) line represents a column with free variables. The remaining columns are pivot columns.

Since a column of the RREF of $\boldsymbol{A}$ is either a pivot column, or has a free variable (see Figure 1.2), upon using the fact that sum of the number of pivot columns and the number of free variables is the total number of columns, we get:

Lemma 1.7.8. *For the matrix* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$*,*

$$\mathrm{rank}(\boldsymbol{A}) + \dim[\mathrm{Null}(\boldsymbol{A})] = n.$$

The dimension of the column space gives us one more bit of information. Suppose that $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, so that $\mathrm{Col}(\boldsymbol{A}) \subset \mathbb{R}^m$. Upon invoking a paraphrase of Theorem 1.5.3, we know that $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for any $\boldsymbol{b} \in \mathbb{R}^m$ if and only if the RREF of $\boldsymbol{A}$ has precisely m pivot columns. In other words, the system is consistent for any $\boldsymbol{b}$ if and only if

$$\mathrm{rank}(\boldsymbol{A}) = \dim[\mathbb{R}^m] = m \quad \Leftrightarrow \quad \dim[\mathrm{Null}(\boldsymbol{A})] = n - m.$$

If $\mathrm{rank}(\boldsymbol{A}) \leq m-1$, then it will necessarily be the case that $\boldsymbol{Ax} = \boldsymbol{b}$ will not be consistent for all $\boldsymbol{b}$. For example, if $\boldsymbol{A} \in \mathbb{R}^{3\times 3}$ and $\mathrm{rank}(\boldsymbol{A}) = 2$, then it will be the case that the subspace $\mathrm{Col}(\boldsymbol{A})$ is a plane, and the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ will be consistent if and only if the vector $\boldsymbol{b}$ is parallel to that plane.

We now restate the equivalency Theorem 1.5.3 and equivalency Theorem 1.5.5. The first theorem discusses conditions which ensure that a linear system always has a unique solution:

If any of the equivalent conditions Theorem 1.7.9(a)-(c) hold, then $\boldsymbol{A}$ is said to have full rank.

Theorem 1.7.9. *The following statements about a matrix* $\mathbb{R}^{m\times n}$ *are equivalent:*

(a) *the linear system* $\boldsymbol{Ax} = \boldsymbol{b}$ *is consistent for any* $\boldsymbol{b}$
(b) $\dim[\mathrm{Null}(\boldsymbol{A})] = n - m$
(c) $\mathrm{rank}(\boldsymbol{A}) = m$.

This next result gives conditions which ensure that consistent systems have unique solutions:

Theorem 1.7.10. *The following statements about a matrix $\mathbb{R}^{m\times n}$ are equivalent:*

(a) *there is at most one solution to the linear system* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$
(b) $\dim[\mathrm{Null}(\boldsymbol{A})] = 0$
(c) $\mathrm{rank}(\boldsymbol{A}) = n$.

If we wish that the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ be both consistent for all $\boldsymbol{b}$, and to have only unique solutions, then we saw in Chapter 1.5.4 that this is possible only if $\boldsymbol{A}$ is a square matrix, i.e., $m = n$. If the solution is to be unique, then by Theorem 1.7.9(b) we must have

$$\dim[\mathrm{Null}(\boldsymbol{A})] = 0.$$

If the linear system is to be consistent, then by Theorem 1.7.10(c) the rank of the matrix must be the number of rows, i.e.,

$$\mathrm{rank}(\boldsymbol{A}) = n.$$

In terms of dimensions we can then restate Theorem 1.5.9 to say:

Theorem 1.7.11. *Consider the linear system* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$*, where* $\boldsymbol{A} \in \mathbb{R}^{n\times n}$*. The following statements are equivalent:*

(a) *there is a unique solution to the linear system for any* $\boldsymbol{b}$
(b) $\mathrm{rank}(\boldsymbol{A}) = n$
(c) $\dim[\mathrm{Null}(\boldsymbol{A})] = 0$.

◁ *Example 1.7.12.* Suppose that

$$\boldsymbol{A} = \begin{pmatrix} 1 & 3 & -2 & 1 \\ 1 & -1 & 2 & 1 \\ 3 & 4 & -1 & 1 \end{pmatrix}.$$

The RREF of $\boldsymbol{A}$ is given by

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The reduced linear system corresponding to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ is then

$$x_1 + x_3 = 0, \quad x_2 - x_3 = 0, \quad x_4 = 0,$$

so that

$$\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}\}.$$

Since the pivot columns are the first, second, and fourth columns of the RREF of $\boldsymbol{A}$, a basis for $\mathrm{Col}(\boldsymbol{A})$ is given by the set

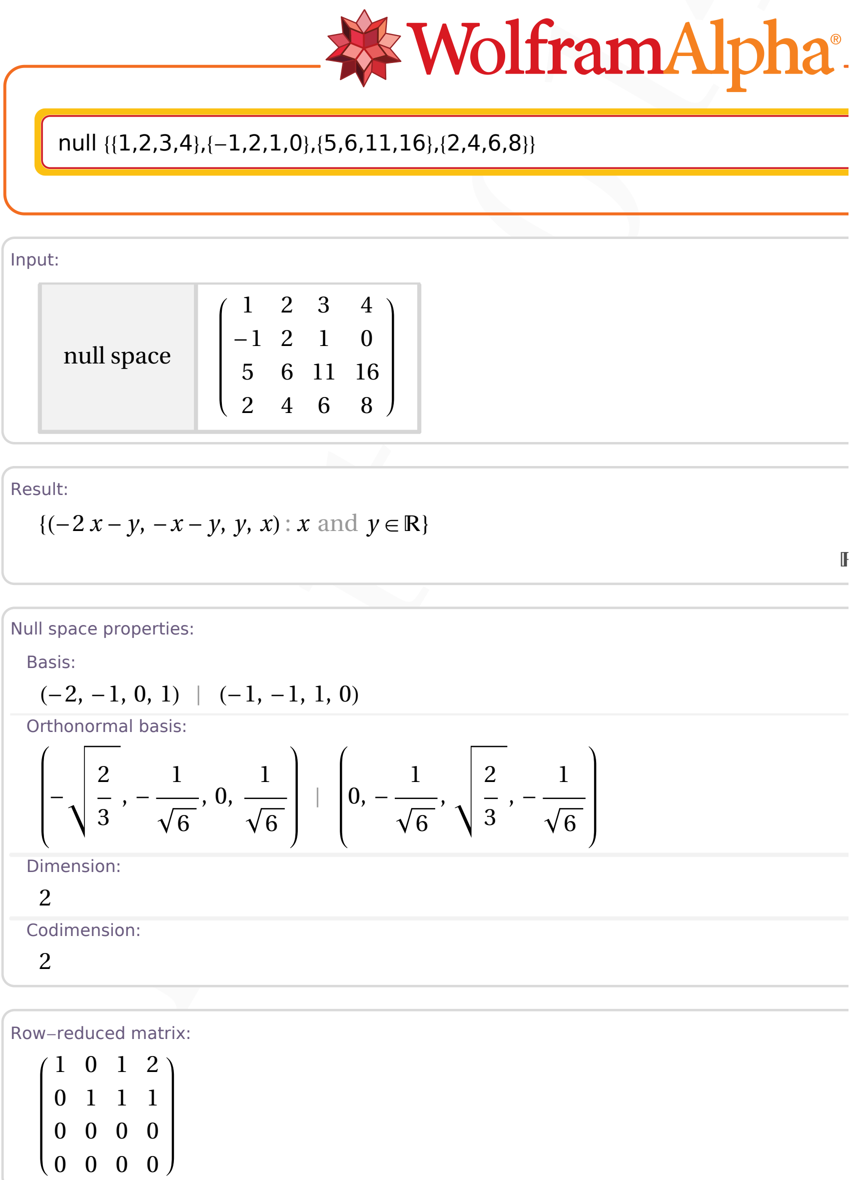
$$\{\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\}.$$

Since $\text{rank}(\boldsymbol{A}) = 3 = \dim[\mathbb{R}^3]$, the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent for all $\boldsymbol{x} \in \mathbb{R}^3$. Since $\dim[\text{Null}(\boldsymbol{A})] = 1 > 0$, the solutions are not unique.

◁ *Example* 1.7.13. We now show how $\text{Col}(\boldsymbol{A})$ and $\text{Null}(\boldsymbol{A})$ are computed using WolframAlpha. Here $\boldsymbol{A} \in \mathbb{R}^{4\times 4}$ is given by

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & 2 & 1 & 0 \\ 5 & 6 & 11 & 16 \\ 2 & 4 & 6 & 8 \end{pmatrix}.$$

We start with



The RREF of $\boldsymbol{A}$ is given in the bottom panel, and from that we see $\mathrm{rank}(\boldsymbol{A}) = 2$, and $\dim[\mathrm{Null}(\boldsymbol{A})] = 2$. A basis for $\mathrm{Col}(\boldsymbol{A})$ is the pivot columns of $\boldsymbol{A}$, which are the first two columns, so

$$\mathrm{Col}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} 1 \\ -1 \\ 5 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 6 \\ 4 \end{pmatrix}\}.$$

As for the null space we read

$$\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} -2 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}\}.$$

Exercises

Exercise 1.7.1. For each of the following matrices not only find a basis for $\mathrm{Col}(\boldsymbol{A})$ and $\mathrm{Null}(\boldsymbol{A})$, but determine $\mathrm{rank}(\boldsymbol{A})$ and $\dim[\mathrm{Null}(\boldsymbol{A})]$.

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 4 \\ 4 & 7 & 8 \end{pmatrix}$

(b) $\boldsymbol{A} = \begin{pmatrix} -3 & 1 & 3 & 4 \\ 1 & 2 & -1 & -2 \\ -3 & 8 & 4 & 2 \end{pmatrix}$

(c) $\boldsymbol{A} = \begin{pmatrix} 1 & 3 & -2 & 1 \\ 2 & 1 & 3 & 2 \\ 3 & 4 & 5 & 6 \end{pmatrix}$

Exercise 1.7.2. Set

$$\boldsymbol{A} = \begin{pmatrix} 1 & -3 & -3 \\ 2 & 5 & -4 \end{pmatrix}, \quad \boldsymbol{v} = \begin{pmatrix} 2 \\ -4 \\ 7 \end{pmatrix}.$$

(a) Is $\boldsymbol{v} \in \mathrm{Null}(\boldsymbol{A})$? Explain.
(b) Describe all vectors that belong to $\mathrm{Null}(\boldsymbol{A})$ as the span of a finite set of vectors.
(c) What is $\dim[\mathrm{Null}(\boldsymbol{A})]$?

Exercise 1.7.3. Set

$$\boldsymbol{A} = \begin{pmatrix} 1 & -3 \\ 2 & 5 \\ -1 & 4 \end{pmatrix}, \quad \boldsymbol{u} = \begin{pmatrix} 2 \\ -4 \\ 7 \end{pmatrix}, \quad \boldsymbol{v} = \begin{pmatrix} -3 \\ 16 \\ 5 \end{pmatrix}.$$

(a) Is $\boldsymbol{u} \in \mathrm{Col}(\boldsymbol{A})$? Explain.
(b) Is $\boldsymbol{v} \in \mathrm{Col}(\boldsymbol{A})$? Explain.
(c) Describe all vectors that belong to $\mathrm{Col}(\boldsymbol{A})$ as the span of a finite set of vectors.
(d) What is $\mathrm{rank}(\boldsymbol{A})$?

Exercise 1.7.4. Can a set of eight vectors be a basis for $\mathbb{R}^7$? Explain.

Exercise 1.7.5. Can a set of five vectors be a basis for $\mathbb{R}^6$? Explain.

Exercise 1.7.6. Is the set

$$S = \{\begin{pmatrix} 1 \\ 5 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \end{pmatrix}\}$$

a basis for $\mathbb{R}^2$? Explain. If not, find a basis for $\mathrm{Span}(S)$, and determine $\dim[\mathrm{Span}(S)]$.

Exercise 1.7.7. Is the set

$$S = \{\begin{pmatrix} 1 \\ 5 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ -5 \end{pmatrix}\}$$

a basis for $\mathbb{R}^3$? Explain. If not, find a basis for $\mathrm{Span}(S)$, and determine $\dim[\mathrm{Span}(S)]$.

Exercise 1.7.8. Is the set

$$S = \{\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}\}$$

a basis for $\mathbb{R}^3$? Explain. If not, find a basis for $\mathrm{Span}(S)$, and determine $\dim[\mathrm{Span}(S)]$.

Exercise 1.7.9. Set

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 5 \\ -1 & 5 & 2 \\ 2 & -7 & -1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 8 \\ 13 \\ -17 \end{pmatrix}.$$

(a) Find a basis for $\mathrm{Col}(\boldsymbol{A})$.
(b) What is $\mathrm{rank}(\boldsymbol{A})$?
(c) The vector $\boldsymbol{b} \in \mathrm{Col}(\boldsymbol{A})$. Write the vector as a linear combination of the basis vectors chosen in part (a).

Exercise 1.7.10. Set

$$\boldsymbol{A} = \begin{pmatrix} 1 & -3 & -2 & 0 \\ 2 & -6 & 1 & 5 \\ -1 & 3 & 3 & 1 \\ -3 & 9 & 1 & -5 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} -5 \\ 5 \\ 8 \\ 0 \end{pmatrix}.$$

(a) Find a basis for $\mathrm{Col}(\boldsymbol{A})$.
(b) What is $\mathrm{rank}(\boldsymbol{A})$?
(c) The vector $\boldsymbol{b} \in \mathrm{Col}(\boldsymbol{A})$. Write the vector as a linear combination of the basis vectors chosen in part (a).

Exercise 1.7.11. Determine if each of the following statements is true or false. Provide an explanation for your answer.

(a) If $\boldsymbol{A} \in \mathbb{R}^{7\times 7}$ is such that the RREF of $\boldsymbol{A}$ has two zero rows, then $\mathrm{rank}(\boldsymbol{A}) = 6$.
(b) Any set of seven linearly independent vectors is a basis for $\mathbb{R}^7$.
(c) If $\boldsymbol{A} \in \mathbb{R}^{4\times 6}$ is such that the RREF of $\boldsymbol{A}$ has one zero row, then $\dim[\mathrm{Null}(\boldsymbol{A})] = 4$.
(d) If $\boldsymbol{A} \in \mathbb{R}^{9\times 9}$ is such that the RREF of $\boldsymbol{A}$ has six pivot columns, $\dim[\mathrm{Null}(\boldsymbol{A})] = 3$.

1.8 Inner-products and orthogonal bases

In Calculus we introduced the *dot product* of two vectors in order to compute the angle between them. We now generalize this notion to vectors of any size, and then discuss the implications of the generalization.

1.8.1 *The inner-product on* $\mathbb{R}^n$

The dot product of two vectors $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^3$ is given by

$$\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{j=1}^{3} x_j y_j = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

The natural generalization to vectors in $\mathbb{R}^n$ is:

Inner-product

Definition 1.8.1. An *inner-product* on $\mathbb{R}^n$ is given by

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{j=1}^{n} x_j y_j = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

The inner-product of two vectors has the same properties as the dot product:

(a) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$
(b) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \ge 0$ with equality if and only if $\boldsymbol{x} = \boldsymbol{0}$
(c) $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$
(d) $\langle c\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, c\boldsymbol{y} \rangle = c\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ for any $c \in \mathbb{R}$

The proof of these properties is left as an exercise in Exercise 1.8.1. Properties (c) and (d) guarantee that the inner product is a linear operation (see Exercise 1.8.2). Property (b) allows us to define the length (magnitude) of a vector by

$$\|\boldsymbol{x}\|^2 := \langle \boldsymbol{x}, \boldsymbol{x} \rangle.$$

The length has the properties

(a) $\|c\boldsymbol{x}\| = |c|\,\|\boldsymbol{x}\|$ for any $c \in \mathbb{R}$
(b) $\|\boldsymbol{x} + \boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}|$ (the triangle inequality)

(see Exercise 1.8.3). A *unit vector* has length one,

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \|\boldsymbol{x}\|^2 = 1.$$

If $\boldsymbol{x}$ is a nonzero vector, then $\boldsymbol{u} = \boldsymbol{x}/\|\boldsymbol{x}\|$ is a unit vector which points in the same direction as $\boldsymbol{x}$ (see Exercise 1.8.4).

As is the case with the dot product, the inner-product can be used to determine the angle between two vectors. We start by using linearity to say

$$\begin{aligned}\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle &= \langle \boldsymbol{x}, \boldsymbol{x} \rangle - \langle \boldsymbol{y}, \boldsymbol{x} \rangle - \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{y}, \boldsymbol{y} \rangle \\ &= \langle \boldsymbol{x}, \boldsymbol{x} \rangle - 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{y}, \boldsymbol{y} \rangle.\end{aligned}$$

The second line follows from property (a). Using the definition of a length of a vector we can rewrite the above as

$$\|\boldsymbol{x} - \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle.$$

Now, if we think of the three vectors $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}$ as forming three legs of a triangle, then by the law of cosines we have

$$\|\boldsymbol{x}-\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - 2\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|\cos\theta,$$

where θ is the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$. Comparing the two equations reveals:

Proposition 1.8.2. *The angle between two vectors $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$ is determined by*

$$\langle \boldsymbol{x},\boldsymbol{y}\rangle = \|\boldsymbol{x}\|\,\|\boldsymbol{y}\|\cos\theta.$$

If the inner-product between two vectors is zero, we say:

Orthogonal

Definition 1.8.3. The two vectors $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$ are *orthogonal* (perpendicular) if

$$\langle \boldsymbol{x},\boldsymbol{y}\rangle = 0.$$

1.8.2 Orthonormal bases and the Gram-Schmidt procedure

Consider a collection of nonzero vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$, and suppose that the vectors are mutually orthogonal, i.e.,

$$\langle \boldsymbol{a}_i, \boldsymbol{a}_j\rangle = 0, \quad i \neq j.$$

We first show that the vectors must be linearly independent. Going back to the original Definition 1.3.5, this means that we must show that the only solution to the homogeneous linear system,

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k = \boldsymbol{0},$$

is $x_1 = x_2 = \cdots = x_k = 0$.

For a given vector $\boldsymbol{y}$ take the inner-product to both sides,

$$\langle x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_k\boldsymbol{a}_k, \boldsymbol{y}\rangle = \langle \boldsymbol{0}, \boldsymbol{y}\rangle.$$

Now use the linearity of the inner-product to rewrite the above as one linear equation,

$$x_1\langle \boldsymbol{a}_1, \boldsymbol{y}\rangle + x_2\langle \boldsymbol{a}_2, \boldsymbol{y}\rangle + \cdots + x_k\langle \boldsymbol{a}_k, \boldsymbol{y}\rangle = 0. \tag{1.8.1}$$

This is one linear equation in k variables; however, we have some freedom in choosing the coefficients by choosing different vectors $\boldsymbol{y}$.

If we choose $\boldsymbol{y} = \boldsymbol{a}_1$ in (1.8.1), then upon using the fact that the vectors are mutually orthogonal the equation reduces to

$$x_1\langle \boldsymbol{a}_1, \boldsymbol{a}_1\rangle = 0.$$

Since the vectors being nonzero implies $\langle \boldsymbol{a}_1, \boldsymbol{a}_1\rangle = \|\boldsymbol{a}_1\|^2 > 0$, we can conclude $x_1 = 0$; thus, (1.8.1) can be rewritten as

$$x_2\langle \boldsymbol{a}_2, \boldsymbol{y}\rangle + x_3\langle \boldsymbol{a}_3, \boldsymbol{y}\rangle + \cdots + x_k\langle \boldsymbol{a}_k, \boldsymbol{y}\rangle = 0. \tag{1.8.2}$$

If we choose $\boldsymbol{y} = \boldsymbol{a}_2$ in (1.8.2), then the mutual orthogonality of the vectors yields the reduced equation

$$x_2\langle \boldsymbol{a}_2, \boldsymbol{a}_2\rangle = 0.$$

Since all of the vectors are nonzero we conclude $x_2 = 0$. Continuing in this fashion leads to $x_3 = x_4 = \cdots = x_k = 0$. We conclude:

Lemma 1.8.4. *Suppose the nonzero vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k\}$ are mutually orthogonal,*

$$\langle \boldsymbol{a}_i, \boldsymbol{a}_j\rangle = 0, \quad i \neq j.$$

The set of vectors is then linearly independent.

We know that while a basis for a subspace is not unique, the dimension is fixed. We now consider the problem of finding an *orthonormal basis* for a given subspace.

Orthonormal basis

Definition 1.8.5. The set of vectors $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k\}$ is *orthonormal* if

$$\langle \boldsymbol{u}_i, \boldsymbol{u}_j\rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

In other words, the vectors are mutually orthogonal, and each has length one.

We now consider the problem of finding an orthonormal basis for a given subspace S. First, suppose that $S = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$, where $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are linearly independent. We wish to find vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ such that:

(a) $S = \mathrm{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$
(b) the set $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ is orthonormal.

We start by setting

$$\boldsymbol{u}_1 = \frac{1}{\|\boldsymbol{a}_1\|}\boldsymbol{a}_1,$$

which is a unit vector pointing in the same direction as $\boldsymbol{a}_1$. Consequently, we can write $S = \mathrm{Span}\{\boldsymbol{u}_1, \boldsymbol{a}_2\}$. In order to satisfy property (a) we first consider a vector, $\boldsymbol{w}_2$, which is a linear combination of $\boldsymbol{u}_1$ and $\boldsymbol{a}_2$,

$$\boldsymbol{w}_2 = \boldsymbol{a}_2 + c_1\boldsymbol{u}_1.$$

It is the case that $S = \mathrm{Span}\{\boldsymbol{u}_1, \boldsymbol{w}_2\}$. Requiring this new vector to be orthogonal to $\boldsymbol{u}_1$ means

$$0 = \langle \boldsymbol{w}_2, \boldsymbol{u}_1\rangle = \langle \boldsymbol{a}_2, \boldsymbol{u}_1\rangle + c_1\langle \boldsymbol{u}_1, \boldsymbol{u}_1\rangle.$$

Since $\boldsymbol{u}_1$ is a unit vector, we have

$$0 = \langle \boldsymbol{a}_2, \boldsymbol{u}_1\rangle + c_1 \quad \Rightarrow \quad c_1 = -\langle \boldsymbol{a}_2, \boldsymbol{u}_1\rangle.$$

The vector

$$\boldsymbol{w}_2 = \boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{u}_1\rangle\boldsymbol{u}_1$$

is orthogonal to $\boldsymbol{u}_1$. Upon normalizing the vector $\boldsymbol{w}_2$ (scaling it to have length one) we have $S = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, where

$$\begin{aligned} \boldsymbol{u}_1 &= \frac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|} \\ \boldsymbol{u}_2 &= \frac{\boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1}{\|\boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1\|}. \end{aligned} \tag{1.8.3}$$

By Lemma 1.8.4 the set $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ is linearly independent; consequently, they form an orthonormal basis for S.

◁ *Example* *1.8.6.* Suppose that $S = \text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$, where

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 2 \\ 7 \\ 4 \end{pmatrix}.$$

Since $\|\boldsymbol{a}_1\|^2 = \langle \boldsymbol{a}_1, \boldsymbol{a}_1 \rangle = 14$, we see from (1.8.3) that the first unit vector is

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Now consider the second vector. Since

$$\langle \boldsymbol{a}_2, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 = \frac{1}{14} \langle \boldsymbol{a}_2, \boldsymbol{a}_1 \rangle \boldsymbol{a}_1 = \frac{28}{14} \boldsymbol{a}_1,$$

we have

$$\boldsymbol{w}_2 = \boldsymbol{a}_2 - 2\boldsymbol{a}_1 = \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}.$$

The second unit vector is

$$\boldsymbol{u}_2 = \frac{\boldsymbol{w}_2}{\|\boldsymbol{w}_2\|} = \frac{1}{\sqrt{13}} \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}.$$

The set $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ is an orthonormal basis for S.

Now suppose that $S = \text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$, where the set of vectors is linearly independent. We wish to find vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ such that:

(a) $S = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$
(b) the set $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is orthonormal.

We have $\text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\} = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, where the orthonormal vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are givn in (1.8.3). Thus, all that is needed is to find the third vector $\boldsymbol{u}_3$. We start with a vector that is a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{a}_3$,

$$\boldsymbol{w}_3 = \boldsymbol{a}_3 + c_1 \boldsymbol{u}_1 + c_2 \boldsymbol{u}_2,$$

so $\text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\} = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{w}_3\}$. We want $\boldsymbol{w}_3$ to be orthogonal to both $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. Requiring that this vector be orthogonal to $\boldsymbol{u}_1$ means

$$0 = \langle \boldsymbol{w}_3, \boldsymbol{u}_1 \rangle = \langle \boldsymbol{a}_3, \boldsymbol{u}_1 \rangle + c_1 \langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle + c_2 \langle \boldsymbol{u}_2, \boldsymbol{u}_1 \rangle.$$

Since $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ is an orthonormal set,

$$\langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle = 1, \quad \langle \boldsymbol{u}_2, \boldsymbol{u}_1 \rangle = 0,$$

and the above equation collapses to

$$0 = \langle \boldsymbol{a}_3, \boldsymbol{u}_1 \rangle + c_1 \quad \Rightarrow \quad c_1 = -\langle \boldsymbol{a}_3, \boldsymbol{u}_1 \rangle.$$

Requiring that $\langle \boldsymbol{w}_3, \boldsymbol{u}_2 \rangle = 0$ and following the same argument gives

$$c_2 = -\langle \boldsymbol{a}_3, \boldsymbol{u}_2 \rangle.$$

The vector

$$\boldsymbol{w}_3 = \boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 - \langle \boldsymbol{a}_3, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2$$

is then orthogonal to both $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. The desired orthonormal set comes upon normalizing $\boldsymbol{w}_3$,

$$\boldsymbol{u}_3 = \frac{\boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 - \langle \boldsymbol{a}_3, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2}{\|\boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 - \langle \boldsymbol{a}_3, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2\|}.$$

We can clearly continue this process, which is known as the *Gram-Schmidt procedure*, for any finite collection of vectors. Doing so yields the following algorithm:

Gram-Schmidt procedure

Lemma 1.8.7. *Let* $S = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k\}$. *An orthonormal basis for* S *is found through the algorithm:*

$$\begin{aligned}
\boldsymbol{u}_1 &= \frac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|} \\
\boldsymbol{u}_2 &= \frac{\boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1}{\|\boldsymbol{a}_2 - \langle \boldsymbol{a}_2, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1\|} \\
\boldsymbol{u}_3 &= \frac{\boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 - \langle \boldsymbol{a}_3, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2}{\|\boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 - \langle \boldsymbol{a}_3, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2\|} \\
&\vdots \\
\boldsymbol{u}_k &= \frac{\boldsymbol{a}_k - \sum_{j=1}^{k} \langle \boldsymbol{a}_k, \boldsymbol{u}_j \rangle \boldsymbol{u}_j}{\|\boldsymbol{a}_k - \sum_{j=1}^{k} \langle \boldsymbol{a}_k, \boldsymbol{u}_j \rangle \boldsymbol{u}_j\|}.
\end{aligned}$$

The Gram-Schmidt procedure does not require that we start with a collection of linearly independent vectors. If the set is linearly dependent, then applying the algorithm will still lead to an orthonormal set of vectors which serve as a basis. All that happens is the total number of vectors is reduced. For example, if we start with 5 vectors, but only 3 of them are linearly independent, then the Gram-Schmidt procedure will lead to a set of 3 orthonormal vectors. The set of orthonormal vectors will be a basis for the span of the original set of 5 vectors.

◁ *Example 1.8.8.* Let us consider a variant of Example 1.8.6. Suppose that $S = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$, where

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 2 \\ 7 \\ 4 \end{pmatrix}, \quad \boldsymbol{a}_3 = \begin{pmatrix} 1 \\ 8 \\ -1 \end{pmatrix}.$$

We have already seen $\text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\} = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, where

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{14}}\begin{pmatrix}1\\2\\3\end{pmatrix}, \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{13}}\begin{pmatrix}0\\3\\-2\end{pmatrix}.$$

For the third vector we start with

$$\boldsymbol{w}_3 = \boldsymbol{a}_3 - \langle \boldsymbol{a}_3, \boldsymbol{u}_1\rangle \boldsymbol{u}_1 - \langle \boldsymbol{a}_3, \boldsymbol{u}_2\rangle \boldsymbol{u}_2.$$

Since

$$\langle \boldsymbol{a}_3, \boldsymbol{u}_1\rangle \boldsymbol{u}_1 = \frac{14}{14}\begin{pmatrix}1\\2\\3\end{pmatrix}, \quad \langle \boldsymbol{a}_3, \boldsymbol{u}_2\rangle \boldsymbol{u}_2 = \frac{26}{13}\begin{pmatrix}0\\3\\-2\end{pmatrix},$$

we have

$$\boldsymbol{w}_3 = \begin{pmatrix}1\\8\\-1\end{pmatrix} - \begin{pmatrix}1\\2\\3\end{pmatrix} - 2\begin{pmatrix}0\\3\\-2\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Consequently, there is no third vector in S which is perpendicular to the first two, so $S = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$. The reason that $\boldsymbol{w}_3 = \boldsymbol{0}$ is that the vector $\boldsymbol{a}_3$ is a linear combination of the first two, $\boldsymbol{a}_3 = -3\boldsymbol{a}_1 + 2\boldsymbol{a}_2$. Consequently, $\dim[S] = 2$, so there can be only two basis vectors.

1.8.3 Orthonormal bases and Fourier expansions

We now show that if a basis for a subspace is an orthonormal set of vectors, then it is straightforward to compute the weights associated with any vector which is some linear combination of the orthonormal vectors. Suppose that $S = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k\}$, where $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k\}$ is an orthonormal set of vectors. By Lemma 1.8.4 the set of vectors is linearly independent, so B is a basis for S. If we originally had $S = \text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_\ell\}$ for some $\ell \geq k$, then we know that we can derive the basis B by using the Gram-Schmidt procedure outlined in Lemma 1.8.7.

Now suppose that $\boldsymbol{b} \in S$. There then exist weights $x_1, x_2, \dots, x_k$ such that

$$\boldsymbol{b} = x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + \cdots + x_k\boldsymbol{u}_k.$$

The weights are typically found by solving the linear system

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}, \quad \boldsymbol{A} = (\boldsymbol{u}_1\, \boldsymbol{u}_2\, \dots\, \boldsymbol{u}_n).$$

Since the basis B is an orthonormal set of vectors, we have a relatively straightforward solution to this system. Moreover, there is a relationship between the length of $\boldsymbol{b}$ and the size of the weights. As a consequence of the formulation of the solution, and the subsequent relationship to Fourier series (which is not at all obvious at this point in time), we will call the weights the *Fourier coefficients*, and we will call the representation in terms of the eigenvectors a *Fourier expansion*.

Lemma 1.8.9. *Let* $S = \mathrm{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$*, where* $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ *is an orthonormal set of vectors. If* $\boldsymbol{b} \in S$*, then there is the Fourier expansion*

$$\boldsymbol{b} = \langle \boldsymbol{b}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \langle \boldsymbol{b}, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2 + \cdots + \langle \boldsymbol{b}, \boldsymbol{u}_k \rangle \boldsymbol{u}_k,$$

and the Fourier coefficients are $\langle \boldsymbol{b}, \boldsymbol{u}_j \rangle$ *for* $j = 1, \ldots, k$*. Moreover, we have a version of Parseval's equality,*

$$\langle \boldsymbol{b}, \boldsymbol{b} \rangle = \langle \boldsymbol{b}, \boldsymbol{u}_1 \rangle^2 + \langle \boldsymbol{b}, \boldsymbol{u}_2 \rangle^2 + \cdots + \langle \boldsymbol{b}, \boldsymbol{u}_k \rangle^2.$$

Proof. In order to prove the expansion result we use the same trick as in the proof of Lemma 1.8.4. Since $\boldsymbol{b} \in S$, we have the linear system

$$\boldsymbol{b} = x_1 \boldsymbol{u}_1 + x_2 \boldsymbol{u}_2 + \cdots + x_k \boldsymbol{u}_k.$$

Upon choosing a vector $\boldsymbol{y}$, taking the inner-product of both sides with respect to this vector, and using the linearity of the inner-product, we can collapse the system to the single linear equation,

$$\langle \boldsymbol{b}, \boldsymbol{y} \rangle = x_1 \langle \boldsymbol{u}_1, \boldsymbol{y} \rangle + x_2 \langle \boldsymbol{u}_2, \boldsymbol{y} \rangle + \cdots + x_k \langle \boldsymbol{u}_k, \boldsymbol{y} \rangle.$$

If we take $\boldsymbol{y} = \boldsymbol{u}_1$, then upon using the fact that the vectors are orthonormal,

$$\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j, \end{cases}$$

the equation becomes

$$\langle \boldsymbol{b}, \boldsymbol{u}_1 \rangle = x_1 \cdot 1.$$

If we take $\boldsymbol{y} = \boldsymbol{u}_2$, then the equation becomes

$$\langle \boldsymbol{b}, \boldsymbol{u}_2 \rangle = x_2 \cdot 1.$$

Continuing in this fashion leads to the desired result,

$$x_j = \langle \boldsymbol{b}, \boldsymbol{u}_j \rangle, \quad j = 1, \ldots, k.$$

Regarding Parseval's equality, this follows immediately upon taking the inner product of both sides of the Fourier expansion with $\boldsymbol{b}$, and using the linearity of the inner product. The details are left for the interested student (see Exercise 1.8.5). □

◁ *Example 1.8.10.* Suppose $S = \mathrm{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, where

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{13}} \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}.$$

We have already seen in Example 1.8.6 that these two vectors form an orthonormal basis for S. Now suppose

$$\boldsymbol{b} = \begin{pmatrix} 1 \\ 8 \\ -1 \end{pmatrix}.$$

The Fourier expansion for $\boldsymbol{b}$ is

$$\boldsymbol{b} = \langle \boldsymbol{b}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \langle \boldsymbol{b}, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2.$$

The Fourier coefficients are

$$\langle \boldsymbol{b}, \boldsymbol{u}_1 \rangle = \frac{14}{\sqrt{14}} = \sqrt{14}, \quad \langle \boldsymbol{b}, \boldsymbol{u}_2 \rangle = \frac{26}{\sqrt{13}} = 2\sqrt{13},$$

so

$$\boldsymbol{b} = \sqrt{14}\,\boldsymbol{u}_1 + 2\sqrt{13}\,\boldsymbol{u}_2.$$

By Parseval's equality the square of the length of $\boldsymbol{b}$ is the sum of the square of the Fourier coefficients,

$$\|\boldsymbol{b}\|^2 = \langle \boldsymbol{b}, \boldsymbol{b} \rangle = (\sqrt{14})^2 + (2\sqrt{13})^2 = 66.$$

1.8.4 Fourier expansions with trigonometric functions

While we have defined an inner-product only for vectors in $\mathbb{R}^n$, the idea can be used in a much more general way. For a concrete example, consider the space of continuous 2π-periodic real-valued functions,

$$C^0_{\text{per}} := \{f : f(x + 2\pi) = f(x),\ f(x) \text{ is continuous}\}. \tag{1.8.4}$$

It can be shown that C^0_{per} is a vector space (see Exercise 1.8.11). Functions in this space include $\cos(x)$, $\sin(3x)$ and $\cos^4(3x)\sin^2(5x)$. The space is important when considering solutions to partial differential equations (see Asmar [4], Haberman [21]), in signal processing (see Oppenheim et al. [35]), and in many other applications.

We define an inner-product on C^0_{per} by

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x)\,\mathrm{d}x.$$

It is straightforward to show that this inner-product has the same properties as that of the inner-product on $\mathbb{R}^n$ (see Exercise 1.8.12). We defined the length of vectors in $\mathbb{R}^n$ through the inner-product on $\mathbb{R}^n$; hence, we can do the same on the space C^0_{per},

$$\|f\|^2 := \langle f, f \rangle.$$

Because the length is defined via the inner-product, it will again have the properties

(a) $\|cf\| = |c|\,\|f\|$ for any $c \in \mathbb{R}$
(b) $\|f + g\| \le \|f\| + \|g\|$ (the triangle inequality)

(see Exercise 1.8.13).

If we have an orthonormal set of functions, then the Fourier expansion result of Lemma 1.8.9 still holds, as this result only depends upon the fact that the set of basis vectors is orthonormal under *some* inner-product. A standard set of orthonormal functions on C^0_{per} is given by $B_N := B_N^{\mathrm{c}} \cup B_N^{\mathrm{s}}$, where

$$B_N^{\mathrm{c}} := \{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos(x), \frac{1}{\sqrt{\pi}}\cos(2x), \dots, \frac{1}{\sqrt{\pi}}\cos(Nx)\}$$
$$B_N^{\mathrm{s}} := \{\frac{1}{\sqrt{\pi}}\sin(x), \frac{1}{\sqrt{\pi}}\sin(2x), \dots, \frac{1}{\sqrt{\pi}}\sin(Nx)\} \tag{1.8.5}$$

(see Exercise 1.8.14). Here $N \geq 1$ is an arbitrary integer. The set B_N is a basis for a subspace of C^0_{per}, and a basis for the full space is achieved upon letting $N \to +\infty$; in other words, $\dim[C^0_{\mathrm{per}}] = \infty$! The verification that one can indeed take the limit is beyond the scope of this text; however, it can be found in Haberman [21]. If $f \in \mathrm{Span}\{B_N\}$, then it will have the expansion

$$f(x) = \frac{1}{2\pi}\langle f, 1\rangle + \frac{1}{\pi}\sum_{j=1}^{N}\langle f, \cos(jx)\rangle + \frac{1}{\pi}\sum_{j=1}^{N}\langle f, \sin(jx)\rangle. \tag{1.8.6}$$

The form of the individual terms follows from some algebraic manipulation, e.g.,

$$\langle f, \frac{1}{\sqrt{\pi}}\cos(jx)\rangle \frac{1}{\sqrt{\pi}}\cos(jx)\rangle = \frac{1}{\pi}\langle f, \cos(jx)\rangle.$$

The term

$$\overline{f} := \frac{1}{2\pi}\langle f, 1\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{d}x$$

is the average (mean) of the function $f(x)$.

◁ *Example 1.8.11.* Let us find the Fourier expansion in B_4 for $f(x) = \sin(x)\sin(2x)$. Since $f(x)$ is an even function, it will be the case that

$$\langle f, \sin(jx)\rangle = 0, \quad j = 1, \dots, 4.$$

A sequence of calculations using WolframAlpha reveals

$$\langle f, 1\rangle = \langle f, \cos(2x)\rangle = \langle f, \cos(4x)\rangle = 0,$$

and

$$\langle f, \cos(x)\rangle = \frac{\pi}{2}, \quad \langle f, \cos(3x)\rangle = -\frac{\pi}{2}.$$

Using the expansion in (1.8.6) with $N = 4$ gives the trigonometric identity

$$\sin(x)\sin(2x) = f(x) = \frac{1}{2}\cos(x) - \frac{1}{2}\cos(3x).$$

Fourier expansions using trigonometric functions have great utility in a wide variety of applications. For example, in the context of signal processing suppose there is a periodic signal. Further suppose that this signal is represented by the function $f(\theta)$, and the periodicity implies $f(\theta) = f(\theta + 2\pi)$. If the signal is continuous, then it can be represented through the *Fourier series*,

$$f(\theta) = \overline{f} + \sum_{j=1}^{\infty} a_j \cos(j\theta) + \sum_{j=1}^{\infty} b_j \sin(j\theta).$$

The series representation is found by taking the limit $N \to +\infty$ in (1.8.6). The Fourier coefficients are given by

$$\overline{f} = \frac{1}{2\pi} \int_0^{2\pi} f(\theta)\,\mathrm{d}\theta, \quad a_j = \frac{1}{\pi} \int_0^{2\pi} f(\theta)\cos(j\theta)\,\mathrm{d}\theta, \quad b_j = \frac{1}{\pi} \int_0^{2\pi} f(\theta)\sin(j\theta)\,\mathrm{d}\theta.$$

Via this series representation we can then think of the signal as the linear superposition of an infinite number of base frequencies.

When studying the signal it is often advantageous to consider the behavior of the Fourier coefficients, as it is not unusual for it to be the case that only a small number of the coefficients are not very small. If this is the case, then we can approximate the signal to good accuracy with the superposition of a relatively small number of frequencies. For example, suppose that for a given signal it is the case that $|a_j| \le 10^{-6}$ for all j, and further suppose that the same upper bound holds for all the b_j's except when $j = 1, 4$. A good approximate representation of the signal would then be

$$f(\theta) \sim \overline{f} + b_1 \sin(\theta) + b_4 \sin(4\theta).$$

The interested reader can consult Haberman [21] for more information, especially in the context of using Fourier series to solve partial differential equations.

Exercises

Exercise 1.8.1. For the inner-product as defined in Definition 1.8.1 show that:

(a) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$
(b) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \ge 0$ with equality if and only if $\boldsymbol{x} = \boldsymbol{0}$
(c) $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$
(d) $\langle c\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, c\boldsymbol{y} \rangle = c\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ for any $c \in \mathbb{R}$.

Exercise 1.8.2. Show that the inner-product as defined in Definition 1.8.1 has the property of linearity,

$$\langle a\boldsymbol{x} + b\boldsymbol{y}, \boldsymbol{z} \rangle = a\langle \boldsymbol{x}, \boldsymbol{z} \rangle + b\langle \boldsymbol{y}, \boldsymbol{z} \rangle.$$

Exercise 1.8.3. The length of vectors in $\mathbb{R}^n$ is defined as

$$\|\boldsymbol{x}\|^2 := \langle \boldsymbol{x}, \boldsymbol{x} \rangle.$$

Show that:

(a) $\|c\boldsymbol{x}\| = |c|\,\|\boldsymbol{x}\|$ for any $c \in \mathbb{R}$
(b) $\|\boldsymbol{x} + \boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}|$ (the triangle inequality)

Exercise 1.8.4. If $\boldsymbol{x} \in \mathbb{R}^n$ is a nonzero vector, show that

$$\boldsymbol{u} = \frac{1}{\|\boldsymbol{x}\|}\boldsymbol{x}$$

is a unit vector.

Exercise 1.8.5. Prove Parseval's equality in Lemma 1.8.9 when:

(a) $k = 2$
(b) $k = 3$
(c) $k \geq 4$ (*Hint*: use an induction argument).

Exercise 1.8.6. Find the angle between the following vectors:

(a) $\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \end{pmatrix}$
(b) $\begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ -4 \\ 7 \end{pmatrix}$
(c) $\begin{pmatrix} 1 \\ 2 \\ 4 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \\ 4 \\ 5 \end{pmatrix}$

Exercise 1.8.7. Find an orthonormal basis for $S = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$, where

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \quad \boldsymbol{a}_3 = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}.$$

Exercise 1.8.8. Find an orthonormal basis for $S = \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$, where

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ -1 \\ 4 \\ 7 \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} -2 \\ 3 \\ -1 \\ 5 \end{pmatrix}.$$

Exercise 1.8.9. Let

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad \boldsymbol{u}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 4 \\ -3 \\ 5 \end{pmatrix}.$$

(a) Show that $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is an orthonormal basis for $\mathbb{R}^3$.
(b) Find the Fourier coefficients associated with the vector $\boldsymbol{b}$.
(c) Find the Fourier expansion for the vector $\boldsymbol{b}$ in terms of the given basis vectors.
(d) Use Parseval's equality to find the length of $\boldsymbol{b}$.

Exercise 1.8.10. Let

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \quad \boldsymbol{u}_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 0 \\ 0 \\ 2 \\ -1 \end{pmatrix}, \quad \boldsymbol{u}_3 = \frac{1}{\sqrt{31}}\begin{pmatrix} -5 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 7 \\ -1 \\ 7 \\ -1 \end{pmatrix}.$$

(a) Show that $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is an orthonormal set of vectors.
(b) Find the Fourier coefficients associated with the vector $\boldsymbol{b}$.
(c) Find the Fourier expansion for the vector $\boldsymbol{b}$ in terms of the given vectors.
(d) Use Parseval's equality to find the length of $\boldsymbol{b}$.

Exercise 1.8.11. Show that the space of continuous 2π-periodic functions, C^0_{per}, as defined in (1.8.4) is a vector space under the definition given in Chapter 1.6.

Exercise 1.8.12. Show that the inner-product on the space C^0_{per} defined by

$$\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\,\mathrm{d}x$$

has the properties:

(a) $\langle f, g\rangle = \langle g, f\rangle$
(b) $\langle f, f\rangle \geq 0$ with equality if and only if $f(x) \equiv 0$
(c) $\langle f+g, h\rangle = \langle f, h\rangle + \langle g, h\rangle$
(d) $\langle cf, g\rangle = \langle f, cg\rangle = c\langle f, g\rangle$ for any $c \in \mathbb{R}$.

Exercise 1.8.13. The length of a function in C^0_{per} is defined by

$$\|f\|^2 = \langle f, f\rangle.$$

Show that:

(a) $\|cf\| = |c|\|f\|$ for any $c \in \mathbb{R}$
(b) $\|f+g\| \leq \|f\| + \|g\|$ (the triangle inequality).

Exercise 1.8.14. Consider the inner-product on the space C^0_{per} defined by

$$\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\,\mathrm{d}x.$$

Show that for any integers $j, k \geq 1$:

(a) $\langle \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{2\pi}}\rangle = 1$
(b) $\langle \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos(jx)\rangle = 0$
(c) $\langle \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\sin(jx)\rangle = 0$
(d) $\langle \frac{1}{\sqrt{\pi}}\cos(jx), \frac{1}{\sqrt{\pi}}\cos(kx)\rangle = \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases}$
(e) $\langle \frac{1}{\sqrt{\pi}}\sin(jx), \frac{1}{\sqrt{\pi}}\sin(kx)\rangle = \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases}$
(f) $\langle \frac{1}{\sqrt{\pi}}\sin(jx), \frac{1}{\sqrt{\pi}}\cos(kx)\rangle = 0.$

Exercise 1.8.15. Find Fourier expansions for the following products of trigonometric functions using the basis B_3, which is defined in (1.8.5):

(a) $\sin(x)\cos(2x)$
(b) $\sin(x)\sin(2x)$
(c) $\cos(x)\cos(2x)$

Exercise 1.8.16. Find Fourier expansions for the following products of trigonometric functions using the basis B_5, which is defined in (1.8.5):

(a) $\sin(x)\sin(2x)\cos(2x)$
(b) $\cos(x)\sin(2x)\cos(2x)$

(c) $\sin(x)\cos(x)\sin(3x)$
(d) $\cos(2x)\cos(3x)$
(e) $\sin(2x)\cos(3x)$
(f) $\sin(2x)\sin(3x)$

1.9 Matrix manipulations

1.9.1 Addition, subtraction, and multiplication

Our goal here is to now consider the algebra of matrices. We will first consider addition, subtraction, and multiplication. Division will be discussed in the next section. The addition and subtraction are straightforward, as is scalar multiplication. If we denote two matrices as $\boldsymbol{A} = (a_{jk}) \in \mathbb{R}^{m\times n}$ and $\boldsymbol{B} = (b_{jk}) \in \mathbb{R}^{m\times n}$, then it is the case that

$$\boldsymbol{A} \pm \boldsymbol{B} = (a_{jk} \pm b_{jk}), \quad c\boldsymbol{A} = (ca_{jk}).$$

In other words, we add/subtract two matrices of the same size component-by-component, and if we multiply a matrix by a scalar, then we multiply each component by that scalar. This is exactly what we do in the addition/subtraction of vectors, and the multiplication of a vector by a scalar. For example, if

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix},$$

then

$$\boldsymbol{A} + \boldsymbol{B} = \begin{pmatrix} 3 & 3 \\ 3 & 0 \end{pmatrix}, \quad 3\boldsymbol{A} = \begin{pmatrix} 3 & 6 \\ -3 & -9 \end{pmatrix}.$$

Regarding the multiplication of two matrices, we simply generalize the matrix/vector multiplication. For a given $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, recall that for $\boldsymbol{b} \in \mathbb{R}^n$,

$$\boldsymbol{A}\boldsymbol{b} = b_1\boldsymbol{a}_1 + b_2\boldsymbol{a}_2 + \cdots + b_n\boldsymbol{a}_n, \quad \boldsymbol{A} = (\boldsymbol{a}_1\ \boldsymbol{a}_2 \cdots \boldsymbol{a}_n).$$

If $\boldsymbol{B} = (\boldsymbol{b}_1\ \boldsymbol{b}_2 \cdots \boldsymbol{b}_\ell) \in \mathbb{R}^{n\times\ell}$ (note that each column $\boldsymbol{b}_j \in \mathbb{R}^n$), we then define the multiplication of $\boldsymbol{A}$ and $\boldsymbol{B}$ by

$$\underbrace{\boldsymbol{A}}_{\mathbb{R}^{m\times n}}\underbrace{\boldsymbol{B}}_{\mathbb{R}^{n\times\ell}} = \underbrace{(\boldsymbol{A}\boldsymbol{b}_1\ \boldsymbol{A}\boldsymbol{b}_2 \cdots \boldsymbol{A}\boldsymbol{b}_\ell)}_{\mathbb{R}^{m\times\ell}}.$$

Note that the number of columns of $\boldsymbol{A}$ must match the number of rows of $\boldsymbol{B}$ in order for the operation to make sense. Further note that the number of rows of the product is the number of rows of $\boldsymbol{A}$, and the number of columns of the product is the number of columns of $\boldsymbol{B}$. For example, if

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ -1 & -3 & 2 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 2 & 1 \\ 4 & 3 \\ 6 & 4 \end{pmatrix},$$

then

$$AB = \left(A\begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} A\begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} \right) = \begin{pmatrix} 28 & 19 \\ -2 & -2 \end{pmatrix} \in \mathbb{R}^{2\times 2},$$

and

$$BA = \left(B\begin{pmatrix} 1 \\ -1 \end{pmatrix} B\begin{pmatrix} 2 \\ -3 \end{pmatrix} B\begin{pmatrix} 3 \\ 2 \end{pmatrix} \right) = \begin{pmatrix} 1 & 1 & 8 \\ 1 & -1 & 18 \\ 2 & 0 & 26 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

As the above example illustrates, it may *not* necessarily be the case that $AB = BA$. In this example changing the order of multiplication leads to a resultant matrices of different sizes. However, even if the resultant matrices are the same size, they need not be the same. Suppose that

$$A = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}.$$

We have

$$AB = \left(A\begin{pmatrix} 2 \\ 4 \end{pmatrix} A\begin{pmatrix} 1 \\ 3 \end{pmatrix} \right) = \begin{pmatrix} 10 & 7 \\ -14 & -10 \end{pmatrix} \in \mathbb{R}^{2\times 2},$$

and

$$BA = \left(B\begin{pmatrix} 1 \\ -1 \end{pmatrix} B\begin{pmatrix} 2 \\ -3 \end{pmatrix} \right) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

These are clearly not the same matrix. Thus, in general we cannot expect matrix multiplication to be commutative.

On the other hand, even though matrix multiplication is not necessarily commutative, it is associative, i.e.,

$$A(B + C) = AB + AC.$$

This fact follows from the fact that matrix/vector multiplication is a linear operation (recall Lemma 1.2.4), and the definition of matrix/matrix multiplication through matrix/vector multiplication. In particular, if we write

$$B = (b_1\, b_2 \cdots b_\ell), \quad C = (c_1\, c_2 \cdots c_\ell),$$

then upon writing

$$B + C = (b_1 + c_1\; b_2 + c_2 \;\cdots\; b_\ell + c_\ell)$$

we have

$$\begin{aligned} A(B + C) &= A(b_1 + c_1\; b_2 + c_2 \cdots b_\ell + c_\ell) \\ &= (A(b_1 + c_1)\; A(b_2 + c_2) \;\cdots\; A(b_\ell + c_\ell)) \\ &= (Ab_1 + Ac_1\; Ab_2 + Ac_2 \;\cdots\; Ab_\ell + Ac_\ell) \\ &= (Ab_1\; Ab_2 \;\cdots\; Ab_\ell) + (Ac_1\; Ac_2 \;\cdots\; Ac_\ell) \\ &= AB + AC. \end{aligned}$$

Indeed, while we will not discuss the details here, it is a fact that just like matrix/vector multiplication, matrix/matrix multiplication is a linear operation,

$$A(bB + cC) = bAB + cAC.$$

There is a special matrix which plays the role of the scalar 1 in matrix multiplication: the identity matrix I_n. If $A \in \mathbb{R}^{m\times n}$, then it is straightforward to check that

$$\boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{A}, \quad \boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A}.$$

In particular, if $\boldsymbol{x} \in \mathbb{R}^n$, then it is true that $\boldsymbol{I}_n\boldsymbol{x} = \boldsymbol{x}$.

As we will later see when we encounter the equation $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ for $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\lambda \in \mathbb{R}$, it will be helpful to rewrite it as

$$\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v} \quad \Rightarrow \quad \boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{I}_n\boldsymbol{v} \quad \Rightarrow \quad (\boldsymbol{A} - \lambda\boldsymbol{I}_n)\boldsymbol{v} = \boldsymbol{0}.$$

In other words, the solution to the original system is realized as an element of the null space of the parameter-dependent matrix $\boldsymbol{A} - \lambda\boldsymbol{I}_n$.

1.9.2 The matrix transpose, and two more subspaces

We can also take the *transpose*, $\boldsymbol{A}^{\mathrm{T}}$, of matrix $\boldsymbol{A}$. Writing $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{m\times n}$ for $i = 1,\dots,m$ and $j = 1,\dots,n$, we have $\boldsymbol{A}^{\mathrm{T}} = (a_{ji}) \in \mathbb{R}^{n\times m}$. In other words, each column of $\boldsymbol{A}$ is a row of $\boldsymbol{A}^{\mathrm{T}}$. For example,

$$\begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix}^{\mathrm{T}} = (-1\;3\;2) \quad \Leftrightarrow \quad (-1\;3\;2)^{\mathrm{T}} = \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix},$$

and

$$\begin{pmatrix} -1 & 3 & 2 \\ 2 & -5 & 8 \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} -1 & 2 \\ 3 & -5 \\ 2 & 8 \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} -1 & 2 \\ 3 & -5 \\ 2 & 8 \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} -1 & 3 & 2 \\ 2 & -5 & 8 \end{pmatrix}.$$

Some properties of the matrix transpose are given in Exercise 1.9.3.

The transpose of a matrix to can be used to "solve" an inconsistent linear system. Start with an inconsistent linear system,

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{A} \in \mathbb{R}^{m\times n}\ (n \le m).$$

Assume that the matrix has full rank, $\mathrm{rank}(\boldsymbol{A}) = n$. If $m < n$, then the matrix cannot have full rank, as the RREF of $\boldsymbol{A}$ must have at least $n - m \ge 1$ free variables.

Such a system will arise, for example, in the context of data fitting. Suppose there are n points, (x_1, y_1), (x_2, y_2), …, (x_n, y_n), and further suppose that each point is thought to lie - but actually does not - on a line of the form $y = a_1 + a_2 x$ for some constants a_1 and a_2. This implies the existence of a linear system,

$$\begin{aligned} a_1 + a_2 x_1 &= y_1 \\ a_1 + a_2 x_2 &= y_2 \\ &\vdots \\ a_1 + a_2 x_n &= y_n, \end{aligned}$$

which in matrix/vector form is

$$\boldsymbol{A}\boldsymbol{a} = \boldsymbol{y}; \quad \boldsymbol{A} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \ \boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \ \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

As long as $x_j \neq x_k$ for $j \neq k$, the matrix $\boldsymbol{A}$ will be of full rank, i.e., $\text{rank}(\boldsymbol{A}) = 2$. It is clear that this overdetermined system will in general be inconsistent, as it is too much to expect that each data point lies on the same line.

Going back to the original system, multiply both sides by $\boldsymbol{A}^{\mathrm{T}}$,

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \quad \Rightarrow \quad \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{x} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b},$$

to get a new linear system. Note that $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is square. It turns out to be the case that the square matrix $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ has full rank (see Exercise 1.9.4), so the new system is consistent and has a unique solution (recall Theorem 1.7.11). These solutions are known as *least-squares solutions*, see Meyer [33, Chapter 5]. While we will not prove it here, the least-squares solution, call it $\boldsymbol{x}_{\ell s}$, minimizes the error,

$$\boldsymbol{x}_{\ell s} = \min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|.$$

The least-squares solution provides the vector in $\text{Col}(\boldsymbol{A})$ which is closest to $\boldsymbol{b}$.

In our discussion of subspaces in Definition 1.4.1 and Definition 1.6.6 we considered two subspaces associated with a matrix: the null space, and the column space. The matrix $\boldsymbol{A}^{\mathrm{T}}$ will also have these two subspaces:

$$\text{Null}(\boldsymbol{A}^{\mathrm{T}}) = \{\boldsymbol{x} : \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{0}\}, \quad \text{Col}(\boldsymbol{A}^{\mathrm{T}}) = \{\boldsymbol{b} : \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{b} \text{ is consistent}\}.$$

Since the columns of $\boldsymbol{A}^{\mathrm{T}}$ are the rows of $\boldsymbol{A}$, the space $\text{Col}(\boldsymbol{A}^{\mathrm{T}})$ is sometimes called the *row space* of $\boldsymbol{A}$.

These four subspaces are intimately related to each other. First note that for $\boldsymbol{A} \in \mathbb{R}^{m\times n}$,

$$\text{Col}(\boldsymbol{A}), \text{Null}(\boldsymbol{A}^{\mathrm{T}}) \subset \mathbb{R}^m, \quad \text{Null}(\boldsymbol{A}), \text{Col}(\boldsymbol{A}^{\mathrm{T}}) \subset \mathbb{R}^n.$$

It can be shown that

$$\boldsymbol{b} \in \text{Col}(\boldsymbol{A}),\ \boldsymbol{x} \in \text{Null}(\boldsymbol{A}^{\mathrm{T}}) \quad \Rightarrow \quad \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = 0, \tag{1.9.1}$$

and

$$\boldsymbol{x} \in \text{Null}(\boldsymbol{A}),\ \boldsymbol{b} \in \text{Col}(\boldsymbol{A}^{\mathrm{T}}) \quad \Rightarrow \quad \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = 0$$

(see Exercise 1.9.10 and Exercise 1.9.11). In Exercise 1.9.9 it is shown that

$$\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = \langle \boldsymbol{b}, \boldsymbol{x} \rangle;$$

hence, this quantity is simply the inner-product. Using the definition of orthogonality provided in Definition 1.8.3, we see that vectors in the column space of $\boldsymbol{A}$ are orthogonal to vectors in the null space of $\boldsymbol{A}^{\mathrm{T}}$.

Armed with the fact of (1.9.1), we can restate the equivalency statement of Theorem 1.6.8(b), which says that the linear system is consistent if and only if $\boldsymbol{b} \in \text{Col}(\boldsymbol{A})$. Instead, we can write that the linear system is consistent if and only if

$$\langle \boldsymbol{b}, \boldsymbol{x} \rangle = 0, \quad \text{for all } \boldsymbol{x} \in \text{Null}(\boldsymbol{A}^{\mathrm{T}}).$$

This latter formulation of consistency is known as the *Fredholm alternative*, and is extremely useful in applications.

Even though the subspaces $\text{Col}(\boldsymbol{A})$ and $\text{Col}(\boldsymbol{A}^{\mathrm{T}})$ may be contained in different spaces, they are related through their dimension. In particular, it can be shown that

$$\text{rank}(\boldsymbol{A}) = \text{rank}(\boldsymbol{A}^{\mathrm{T}})$$

(see Meyer [33, Chapter 4]). Thus, both $\boldsymbol{A}$ and $\boldsymbol{A}^{\mathrm{T}}$ will have the same number of pivot columns, and for $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ the number of free variables for each matrix will be

$$\dim[\text{Null}(\boldsymbol{A})] = n - \text{rank}(\boldsymbol{A}), \quad \dim\left[\text{Null}(\boldsymbol{A}^{\mathrm{T}})\right] = m - \text{rank}(\boldsymbol{A}). \tag{1.9.2}$$

If the matrix is square the two null spaces will have the same dimension; otherwise, they will be of different sizes.

Exercises

Exercise 1.9.1. Let

$$\boldsymbol{A} = \begin{pmatrix} 1 & -2 & -5 \\ -2 & 3 & -7 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 2 & 0 \\ -2 & 3 \\ 1 & 5 \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} 1 & 2 \\ 5 & -4 \\ 3 & -1 \end{pmatrix}.$$

Compute the prescribed algebraic operation if it is well-defined. If it cannot be done, explain why.

(a) $3\boldsymbol{B} - 2\boldsymbol{C}$
(b) $4\boldsymbol{A} + 2\boldsymbol{B}$
(c) $2\boldsymbol{A}^{\mathrm{T}} + 3\boldsymbol{B}$
(d) $\boldsymbol{A} - 2\boldsymbol{C}^{\mathrm{T}}$
(e) $\boldsymbol{AB}$
(f) $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}$
(g) $\boldsymbol{BC}^{\mathrm{T}}$
(h) $(\boldsymbol{BA})^{\mathrm{T}}$
(i) $(2\boldsymbol{A} - \boldsymbol{B}^{\mathrm{T}})^{\mathrm{T}}$

Exercise 1.9.2. Suppose that $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n \times k}$ with $m \neq n$ and $n \neq k$ (i.e., neither matrix is square).

(a) What is the size of $\boldsymbol{AB}$?
(b) Can m, k be chosen so that $\boldsymbol{BA}$ is well-defined? If so, what is the size of $\boldsymbol{BA}$?
(c) Is is possible for $\boldsymbol{AB} = \boldsymbol{BA}$? Explain.

Exercise 1.9.3. Let $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m \times n}$. Regarding properties of the transpose of a matrix, show that:

(a) $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$.
(b) $(c\boldsymbol{A})^{\mathrm{T}} = c\boldsymbol{A}^{\mathrm{T}}$ for any constant $c \in \mathbb{R}$.
(c) $(\boldsymbol{A} + \boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$.
(d) $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

Exercise 1.9.4. Suppose that $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $n \le m$, has full rank. Show that

$$\mathrm{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}) = n.$$

Exercise 1.9.5. Suppose that $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, and set

$$\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}}, \quad \boldsymbol{C} = \boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}.$$

(a) Show that $\boldsymbol{B}$ is *symmetric*, i.e., $\boldsymbol{B}^{\mathrm{T}} = \boldsymbol{B}$ (hint: use Exercise 1.9.3).
(b) Show that $\boldsymbol{C}$ is *skew-symmetric*, i.e., $\boldsymbol{C}^{\mathrm{T}} = -\boldsymbol{C}$ (hint: use Exercise 1.9.3).

Exercise 1.9.6. Show that for any $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ the matrices $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ and $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ are both symmetric.

Exercise 1.9.7. Show that if $\boldsymbol{A}$ is skew-symmetric, then all of its diagonal entries must be zero.

Exercise 1.9.8. Show that any $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ may be written as the sum of a symmetric matrix and a skew-symmetric matrix (hint: use Exercise 1.9.5).

Exercise 1.9.9. Show that the inner-product as defined in Definition 1.8.1 can also be written as

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}.$$

Exercise 1.9.10. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. Show that $\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \rangle$. (*Hint*: Use Exercise 1.9.9 and the fact that $(\boldsymbol{A}\boldsymbol{B})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.)

Exercise 1.9.11. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. If $\boldsymbol{b} \in \mathrm{Col}(\boldsymbol{A})$ and $\boldsymbol{x} \in \mathrm{Null}(\boldsymbol{A}^{\mathrm{T}})$, show that $\langle \boldsymbol{b}, \boldsymbol{x} \rangle = 0$. (*Hint*: Use the definition of $\mathrm{Col}(\boldsymbol{A})$ and Exercise 1.9.10)

Exercise 1.9.12. In Exercise 1.6.3 it was shown that if $A, B \subset \mathbb{R}^m$ are subspaces, then the set,

$$A + B = \{\boldsymbol{x} : \boldsymbol{x} = \boldsymbol{a} + \boldsymbol{b},\ \boldsymbol{a} \in A,\ \boldsymbol{b} \in B\},$$

is also a subspace. Suppose that $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. Show that

$$\mathbb{R}^m = \mathrm{Col}(\boldsymbol{A}) + \mathrm{Null}(\boldsymbol{A}^{\mathrm{T}}).$$

(*Hint*: Use (1.9.2) and Exercise 1.9.11)

Exercise 1.9.13. Suppose that for some $\boldsymbol{A} \in \mathbb{R}^{3\times 5}$,

$$\mathrm{Null}(\boldsymbol{A}^{\mathrm{T}}) = \mathrm{Span}\{\begin{pmatrix} 2 \\ -3 \\ 5 \end{pmatrix}\}.$$

Find a basis for $\mathrm{Col}(\boldsymbol{A})$. (*Hint*: use the Fredholm alternative)

Exercise 1.9.14. Determine if each of the following statements is true or false. Provide an explanation for your answer.

(a) If $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m\times n}$ are the same size, then $\boldsymbol{A}\boldsymbol{B}$ and $\boldsymbol{B}\boldsymbol{A}$ are well-defined.
(b) If $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{5\times 5}$, then $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$.
(c) If $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$ are symmetric, then $(\boldsymbol{A}\boldsymbol{B})^{\mathrm{T}} = \boldsymbol{B}\boldsymbol{A}$.
(d) If $\boldsymbol{A}, \boldsymbol{B}$ are such that $\boldsymbol{A} + \boldsymbol{B}$ is well-defined, then $(\boldsymbol{A} + \boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$.

1.10 The inverse of a square matrix

We know how to do matrix multiplication: how about matrix division? If such a thing exists, then we can (formally) write the solution to linear systems as

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \quad \Rightarrow \quad \boldsymbol{x} = \frac{1}{\boldsymbol{A}}\boldsymbol{b}.$$

Unfortunately, as currently written this calculation makes no sense. However, using the analogy that $1/2$ is the unique number such that $1/2 \cdot 2 = 1$, we could define $1/\boldsymbol{A}$ to be that matrix such that $1/\boldsymbol{A} \cdot \boldsymbol{A} = \boldsymbol{I}_n$. It is not clear that for a given matrix $\boldsymbol{A}$ the corresponding matrix $1/\boldsymbol{A}$ must exist. For an analogy, there is no number $c \in \mathbb{R}$ such that $c \cdot 0 = 1$. Moreover, even if $1/\boldsymbol{A}$ does exist, it is not at all clear as to how it should be computed.

When solving the linear system as above, we are implicitly assuming that

(a) a solution exists for any $\boldsymbol{b}$
(b) the solution is unique.

As we saw in Chapter 1.5.4, these two conditions can be satisfied only if the matrix is square. Consequently, for the rest of the discussion we will consider only square matrices, and we will call $1/\boldsymbol{A}$ the *inverse* of a square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. If it exists, it will be denoted by $\boldsymbol{A}^{-1}$ (think $1/2 = 2^{-1}$), and it will have the property that

$$\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}_n. \tag{1.10.1}$$

Assuming that the inverse exists, it allows us to solve the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ via a matrix/vector multiplication. Namely, we have

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \quad \Rightarrow \quad \boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} \quad \Rightarrow \quad \boldsymbol{I}_n\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} \quad \Rightarrow \quad \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}.$$

Lemma 1.10.1. *Consider the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, where $A \in \mathbb{R}^{n\times n}$ is invertible, i.e., $\boldsymbol{A}^{-1}$ exists. The solution to the linear system is given by*

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}.$$

How do we compute the inverse? Denote $\boldsymbol{A}^{-1} = (\boldsymbol{a}_1^{-1}\,\boldsymbol{a}_2^{-1}\cdots\boldsymbol{a}_n^{-1})$, and let $\boldsymbol{e}_j$ denote the j^{th} column of $\boldsymbol{I}_n$, i.e., $\boldsymbol{I}_n = (\boldsymbol{e}_1\,\boldsymbol{e}_2\cdots\boldsymbol{e}_n)$. Using (1.10.1),

$$\left(\boldsymbol{e}_1\ \boldsymbol{e}_2\ \cdots\ \boldsymbol{e}_n\right) = \boldsymbol{I}_n = \boldsymbol{A}\boldsymbol{A}^{-1} = \left(\boldsymbol{A}\boldsymbol{a}_1^{-1}\ \boldsymbol{A}\boldsymbol{a}_2^{-1}\ \cdots\ \boldsymbol{A}_n^{-1}\right).$$

Equating columns gives $\boldsymbol{A}\boldsymbol{a}_j^{-1} = \boldsymbol{e}_j$ for $j = 1,\dots,n$, so that the j^{th} column of $\boldsymbol{A}^{-1}$ is the solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{e}_j$. From Theorem 1.7.11 it must be the case that if $\boldsymbol{A}^{-1}$ exists, then the RREF of $\boldsymbol{A}$ is $\boldsymbol{I}_n$. This yields for the augmented matrix,

$$(\boldsymbol{A}|\boldsymbol{e}_j) \xrightarrow{\text{RREF}} (\boldsymbol{I}_n|\boldsymbol{a}_j^{-1}), \quad j = 1,\dots,n. \tag{1.10.2}$$

We now consider the collection of linear systems (1.10.2) through a different lens. First consider a general collection of linear systems with the same coefficient matrix,

$$\boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{b}_1,\ \boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{b}_2,\ \dots\ \boldsymbol{A}\boldsymbol{x}_m = \boldsymbol{b}_m.$$

Using the definition of matrix/matrix multiplication, this collection of linear systems can be written more compactly as $\boldsymbol{AX} = \boldsymbol{B}$, where

$$\boldsymbol{X} = (\boldsymbol{x}_1\ \boldsymbol{x}_2\ \cdots\ \boldsymbol{x}_m), \quad \boldsymbol{B} = (\boldsymbol{b}_1\ \boldsymbol{b}_2\ \cdots\ \boldsymbol{b}_m).$$

Solving this new system is accomplished by forming the augmented matrix $(\boldsymbol{A}|\boldsymbol{B})$, and then row-reducing.

Now, (1.10.2) is equivalent to solving n linear systems,

$$\boldsymbol{Ax} = \boldsymbol{e}_j, \quad j = 1,\ldots,n.$$

Using the above, this collection of linear systems can be written more compactly as

$$\boldsymbol{AX} = \boldsymbol{I}_n.$$

Forming the augmented matrix $(\boldsymbol{A}|\boldsymbol{I}_n)$, we find the inverse of $\boldsymbol{A}$ via

$$(\boldsymbol{A}|\boldsymbol{I}_n) \xrightarrow{\text{RREF}} (\boldsymbol{I}_n|\boldsymbol{A}^{-1}).$$

Lemma 1.10.2. *The square matrix* $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ *is invertible if and only if the RREF of* $\boldsymbol{A}$ *is* $\boldsymbol{I}_n$. *The inverse is computed via*

$$(\boldsymbol{A}|\boldsymbol{I}_n) \xrightarrow{\text{RREF}} (\boldsymbol{I}_n|\boldsymbol{A}^{-1}).$$

◁ *Example 1.10.3.* Suppose that

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 2 \\ -6 \end{pmatrix}.$$

We have

$$(\boldsymbol{A}|\boldsymbol{I}_2) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|cc} 1 & 0 & -5 & 2 \\ 0 & 1 & 3 & -1 \end{array}\right) \quad \Rightarrow \quad \boldsymbol{A}^{-1} = \begin{pmatrix} -5 & 2 \\ 3 & -1 \end{pmatrix}.$$

Consequently, the solution to the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is given by

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} = \begin{pmatrix} -22 \\ 12 \end{pmatrix}.$$

◁ *Example 1.10.4.* Suppose that

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}.$$

We have

$$(\boldsymbol{A}|\boldsymbol{I}_2) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & -3 & 1 \end{array}\right);$$

consequently, since the left-hand side of the augmented matrix cannot be row-reduced to $\boldsymbol{I}_2$, $\boldsymbol{A}^{-1}$ does not exist. Since the RREF of $\boldsymbol{A}$ is

$$A \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix},$$

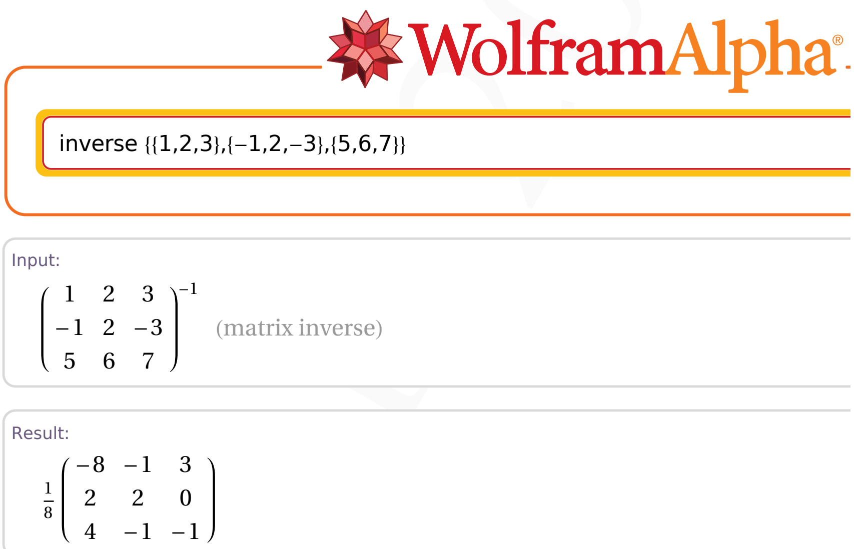
we have that the first column of $\boldsymbol{A}$ is the only pivot column; hence, by Theorem 1.6.8 and the fact that the pivot columns form a basis for $\text{Col}(\boldsymbol{A})$ (see Lemma 1.7.8) the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent if and only if

$$\boldsymbol{b} \in \text{Col}(\boldsymbol{A}) = \text{Span}\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}\}.$$

◁ *Example 1.10.5.* We consider an example for which the inverse will be computed by WolframAlpha. Here $\boldsymbol{A} \in \mathbb{R}^{3\times 3}$ is given by

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & -3 \\ 5 & 6 & 7 \end{pmatrix}.$$

We get



In conclusion,

$$\boldsymbol{A}^{-1} = \frac{1}{8}\begin{pmatrix} -8 & -1 & 3 \\ 2 & 2 & 0 \\ 4 & -1 & -1 \end{pmatrix}.$$

Exercises

Exercise 1.10.1. Find the inverse, if it exists, of the following matrices:

(a) $\begin{pmatrix} 3 & 7 \\ -1 & 4 \end{pmatrix}$

(b) $\begin{pmatrix} -2 & 3 \\ 4 & -6 \end{pmatrix}$

(c) $\begin{pmatrix} 5 & 0 & 0 \\ 0 & 6 & 4 \\ 0 & -2 & -1 \end{pmatrix}$

Exercise 1.10.2. Use A^{-1}, if it exists, to solve the linear system $Ax = b$. If A^{-1} does not exist, find all solutions to the system if it is consistent.

(a) $A = \begin{pmatrix} 3 & 7 \\ -1 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ -6 \end{pmatrix}$
(b) $A = \begin{pmatrix} -2 & 3 \\ 4 & -6 \end{pmatrix}, \quad b = \begin{pmatrix} -4 \\ 8 \end{pmatrix}$
(c) $A = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 6 & 4 \\ 0 & -2 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 1 \\ 9 \end{pmatrix}$

Exercise 1.10.3. Let $A = \begin{pmatrix} 2 & -5 \\ 3 & 4 \end{pmatrix}$, $b_1 = \begin{pmatrix} 7 \\ -8 \end{pmatrix}$, $b_2 = \begin{pmatrix} 0 \\ 6 \end{pmatrix}$. Use A^{-1} to solve the systems $Ax = b_1$ and $Ax = b_2$.

Exercise 1.10.4. Suppose that $A, B \in \mathbb{R}^{n\times n}$ are invertible matrices. Show that

$$(AB)^{-1} = B^{-1}A^{-1}$$

(compare your answer with the result of Exercise 1.9.3(d)).

Exercise 1.10.5. Let $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

(a) Compute A^{-1} and A^{T}. What do you observe?
(b) What is the angle between v and Av?
(c) How does $\|v\|$ compare with $\|Av\|$?

Exercise 1.10.6. Determine if each of the following statements is true or false. Provide an explanation for your answer.

(a) If A has a pivot in every row, then the matrix is invertible.
(b) If $Ax = b$ has a unique solution, then A is invertible.
(c) If $A, B \in \mathbb{R}^{n\times n}$ are invertible, then $(AB)^{-1} = A^{-1}B^{-1}$.
(d) If A is a square matrix whose RREF has one zero row, then A is invertible.
(e) If $A \in \mathbb{R}^{n\times n}$ is invertible, then the columns of A are linearly dependent.
(f) If the RREF of A has no zero rows, then the matrix is invertible.

1.11 The determinant of a square matrix

We wish to derive a scalar which tells us whether or not a square matrix is invertible. First suppose that $A \in \mathbb{R}^{2\times 2}$ is given by

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

If we try to compute A^{-1}, we get

$$(\boldsymbol{A}|\boldsymbol{I}_2) \xrightarrow{-c\rho_1+a\rho_2} \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & ad-bc & -c & a \end{array}\right).$$

If $ad - bc \neq 0$, then we can continue with the row reduction, and eventually compute $\boldsymbol{A}^{-1}$; otherwise, $\boldsymbol{A}^{-1}$ does not exist. This fact implies that this quantity has special significance for 2×2 matrices.

Determinant

Definition 1.11.1. Let $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$ be given by

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The *determinant* of $\boldsymbol{A}$, $\det(\boldsymbol{A})$, is given by

$$\det(\boldsymbol{A}) = ad - bc.$$

We know that the matrix $\boldsymbol{A}$ has the RREF of $\boldsymbol{I}_2$ if and only if $\det(\boldsymbol{A}) \neq 0$. Continuing with the row-reductions if $\det(\boldsymbol{A}) \neq 0$ leads to:

Lemma 1.11.2. *Suppose that* $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$ *is given by*

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The matrix is invertible if and only if $\det(\boldsymbol{A}) \neq 0$. *Furthermore, if* $\det(\boldsymbol{A}) \neq 0$, *then the inverse is given by*

$$\boldsymbol{A}^{-1} = \frac{1}{\det(\boldsymbol{A})} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Proof. A simple calculation shows that $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}_2$ if $\det(\boldsymbol{A}) \neq 0$. □

◁ *Example 1.11.3.* Suppose that

$$\boldsymbol{A} = \begin{pmatrix} 4 & 7 \\ -3 & 2 \end{pmatrix}.$$

Since

$$\det(\boldsymbol{A}) = (4)(2) - (7)(-3) = 29,$$

the inverse of $\boldsymbol{A}$ exists, and it is given by

$$\boldsymbol{A}^{-1} = \frac{1}{29} \begin{pmatrix} 2 & -7 \\ 3 & 4 \end{pmatrix}.$$

By Lemma 1.10.1 the unique solution to the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is given by $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

◁ *Example 1.11.4.* Suppose that

$$\boldsymbol{A} = \begin{pmatrix} 4 & 1 \\ 8 & 2 \end{pmatrix}.$$

Since $\det(\boldsymbol{A}) = 0$, the inverse of $\boldsymbol{A}$ does not exist. If there is a solution to $\boldsymbol{Ax} = \boldsymbol{b}$, it must be found by putting the augmented matrix $(\boldsymbol{A}|\boldsymbol{b})$ into RREF, and then solving the resultant system.

We now wish to define the determinant for $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ for $n \geq 3$. In theory we could derive it in a manner similar to that for the case $n = 2$: start with a matrix of a given size, and then attempt to row-reduce it to the identity. At some point a scalar arises which must be nonzero in order to ensure that the RREF of the matrix is the identity. This scalar would then be denoted as the determinant. Instead of going through this derivation, we instead settle on the final result.

For $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, let $\boldsymbol{A}_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ denote the *submatrix* gotten from $\boldsymbol{A}$ after deleting the i^{th} row and j^{th} column. For example,

$$\boldsymbol{A} = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \quad \Longrightarrow \quad \boldsymbol{A}_{12} = \begin{pmatrix} 2 & 8 \\ 3 & 9 \end{pmatrix}, \quad \boldsymbol{A}_{31} = \begin{pmatrix} 4 & 7 \\ 5 & 8 \end{pmatrix}.$$

With this notion of submatrix in mind, we note that for 2×2 matrices the determinant can be written as

$$\det(\boldsymbol{A}) = a_{11} \det(\boldsymbol{A}_{11}) - a_{12} \det(\boldsymbol{A}_{12}),$$

where here the determinant of a scalar is simply the scalar. The generalization to larger matrices is:

Determinant

Definition 1.11.5. If $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, then the determinant of $\boldsymbol{A}$ is given by

$$\det(\boldsymbol{A}) = a_{11} \det(\boldsymbol{A}_{11}) - a_{12} \det(\boldsymbol{A}_{12}) + a_{13} \det(\boldsymbol{A}_{13}) + \cdots + (-1)^{n+1} a_{1n} \det(\boldsymbol{A}_{1n}).$$

◁ *Example 1.11.6.* If

$$\boldsymbol{A} = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix},$$

then we have

$$\boldsymbol{A}_{11} = \begin{pmatrix} 5 & 8 \\ 6 & 9 \end{pmatrix}, \quad \boldsymbol{A}_{12} = \begin{pmatrix} 2 & 8 \\ 3 & 9 \end{pmatrix}, \quad \boldsymbol{A}_{13} = \begin{pmatrix} 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Since

$$a_{11} = 1, \quad a_{12} = 4, \quad a_{13} = 7,$$

the determinant of the matrix is

$$\det(\boldsymbol{A}) = 1 \cdot \det(\boldsymbol{A}_{11}) - 4 \cdot \det(\boldsymbol{A}_{12}) + 7 \cdot \det(\boldsymbol{A}_{13}) = -3 + 24 - 21 = 0.$$

Thus, we know that $\boldsymbol{A}^{-1}$ does not exist; indeed, the RREF of $\boldsymbol{A}$ is

$$\boldsymbol{A} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

and the columns of $\boldsymbol{A}$ are linearly dependent. Indeed,

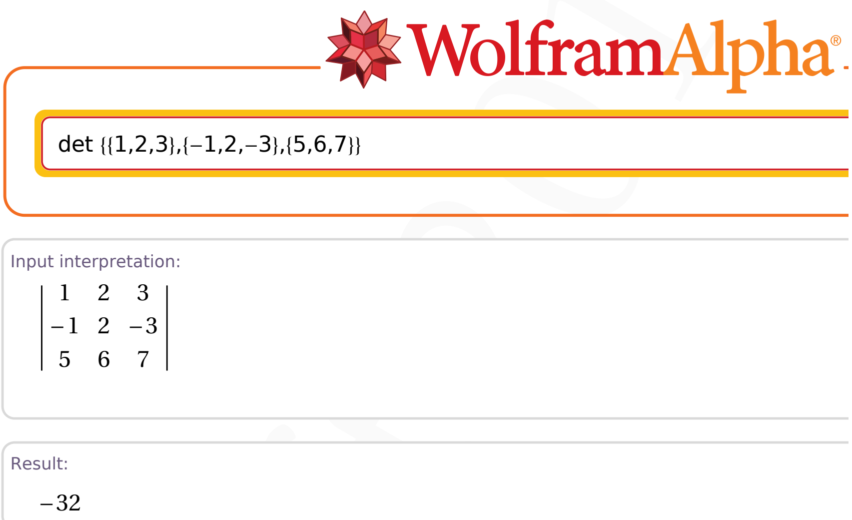
$$\mathrm{Col}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}\}, \quad \mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\},$$

with $\mathrm{rank}(\boldsymbol{A}) = 2$ and $\dim[\mathrm{Null}(\boldsymbol{A})] = 1$. The columns are related through the linear combination $\boldsymbol{a}_1 - 2\boldsymbol{a}_2 + \boldsymbol{a}_3 = \boldsymbol{0}$.

◁ *Example* 1.11.7. We now calculate the determinant using WolframAlpha for the matrix

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & -3 \\ 5 & 6 & 7 \end{pmatrix}.$$

We have



In other words, $\det(\boldsymbol{A}) = -32$.

The determinant has many properties, which are too many to detail in full here (e.g., see Eves [18, Chapter 3] and Vein and Dale [40]). We will consider only a small number that we will directly need. The first, and perhaps most important, is that the expression of Definition 1.11.1 is not the only way to calculate the determinant. In general, the determinant can be calculated by going across any row, or down any column; in particular, we have

$$\det(\boldsymbol{A}) = \underbrace{\sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\boldsymbol{A}_{ij})}_{\text{across } i^{\text{th}} \text{ row}} = \underbrace{\sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(\boldsymbol{A}_{ij})}_{\text{down } j^{\text{th}} \text{ column}}. \tag{1.11.1}$$

For example,

$$\begin{aligned} \det\begin{pmatrix} 4 & 3 & 6 \\ 2 & 0 & 0 \\ -1 & 7 & -5 \end{pmatrix} &= (6)\det\begin{pmatrix} 2 & 0 \\ -1 & 7 \end{pmatrix} - (0)\det\begin{pmatrix} 4 & 3 \\ -1 & 7 \end{pmatrix} + (-5)\det\begin{pmatrix} 4 & 3 \\ 2 & 0 \end{pmatrix} \\ &= -(2)\det\begin{pmatrix} 3 & 6 \\ 7 & -5 \end{pmatrix} + (0)\det\begin{pmatrix} 4 & 6 \\ -1 & -5 \end{pmatrix} - (0)\det\begin{pmatrix} 4 & 3 \\ -1 & 7 \end{pmatrix}. \end{aligned}$$

The first line is down the third column, and the second line is across the second row. As the above example shows, a judicious choice for the expansion of the determinant can greatly simplify the calculation. In particular, it is generally best to calculate the determinant using the row or column which has the most zeros. Note that if a matrix has a zero row or column, then by using the more general definition (1.11.1) and expanding across that zero row or column we get that $\det(\boldsymbol{A}) = 0$.

A couple of other properties which may sometimes be useful are as follows. If a matrix $\boldsymbol{B}$ is formed from $\boldsymbol{A}$ by multiplying a row or column by a constant c, e.g., $\boldsymbol{A} = (\boldsymbol{a}_1\, \boldsymbol{a}_2 \cdots \boldsymbol{a}_n)$ and $\boldsymbol{B} = (c\boldsymbol{a}_1\, \boldsymbol{a}_2 \cdots \boldsymbol{a}_n)$, then $\det(\boldsymbol{B}) = c\det(\boldsymbol{A})$. In particular, after multiplying each column by the same constant, i.e., multiplying the entire matrix by a constant, it is then true that $\det(c\boldsymbol{A}) = c^n \det(\boldsymbol{A})$ (see Exercise 1.11.5). Another useful property is that

$$\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B}).$$

Since $\boldsymbol{I}_n\boldsymbol{A} = \boldsymbol{A}$, we get from this property that

$$\det(\boldsymbol{A}) = \det(\boldsymbol{I}_n)\det(\boldsymbol{A}) = \det(\boldsymbol{I}_n)\det(\boldsymbol{A}) \quad \Rightarrow \quad \det(\boldsymbol{I}_n) = 1$$

(this could also be shown by a direct computation). Since $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}_n$, this also allows us to state that

$$1 = \det(\boldsymbol{I}_n) = \det(\boldsymbol{A}\boldsymbol{A}^{-1}) = \det(\boldsymbol{A})\det(\boldsymbol{A}^{-1}) \quad \Rightarrow \quad \det(\boldsymbol{A}^{-1}) = \frac{1}{\det(\boldsymbol{A})}.$$

We summarize with:

Proposition 1.11.8. *The determinant of matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$ has the properties:*

(a) $\det(c\boldsymbol{A}) = c^n \det(\boldsymbol{A})$
(b) $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$
(c) *if $\boldsymbol{A}$ is invertible*, $\det(\boldsymbol{A}^{-1}) = 1/\det(\boldsymbol{A})$.

Recall that the determinant is defined in such a manner that it provides us with information regarding the solution structure to a linear system of equations. We first list all of the implications of a zero determinant regarding solutions to the linear system $\boldsymbol{Ax} = \boldsymbol{b}$:

Theorem 1.11.9. *Consider the linear system $\boldsymbol{Ax} = \boldsymbol{b}$, where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. The following are equivalent statements:*

(a) $\det(\boldsymbol{A}) = 0$
(b) $\dim[\mathrm{Null}(\boldsymbol{A})] \geq 1$, *i.e., the RREF of $\boldsymbol{A}$ has free variables*
(c) $\mathrm{rank}(\boldsymbol{A}) \leq n - 1$
(d) the inverse matrix $\boldsymbol{A}^{-1}$ does not exist
(e) the RREF of $\boldsymbol{A}$ has at least one zero row
(f) if the linear system is consistent, there are an infinite number of solutions.

On the other hand, if the determinant is nonzero, then we have the following addendum to Theorem 1.7.11:

Theorem 1.11.10. *Consider the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. The following are equivalent:*

(g) $\det(\boldsymbol{A}) \neq 0$
(h) $\dim[\mathrm{Null}(\boldsymbol{A})] = 0$
(i) $\mathrm{rank}(\boldsymbol{A}) = n$, *i.e., $\boldsymbol{A}$ has full rank*
(j) the inverse matrix $\boldsymbol{A}^{-1}$ exists
(k) the RREF of $\boldsymbol{A}$ is the identity matrix $\boldsymbol{I}_n$
(l) the linear system is consistent, and the unique solution is $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

Exercises

Exercise 1.11.1. Compute by hand $\det(\boldsymbol{A})$ for each of the following matrices, and then state whether or not the matrix is invertible. If the matrix is invertible, compute $\det(\boldsymbol{A}^{-1})$.

(a) $\boldsymbol{A} = \begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix}$
(b) $\boldsymbol{A} = \begin{pmatrix} 1 & -3 & 4 \\ 1 & 2 & -1 \\ 3 & -5 & 8 \end{pmatrix}$
(c) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 0 \\ 2 & 8 & 5 \end{pmatrix}$
(d) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 0 & 5 \\ 2 & 4 & 0 & 6 \\ 0 & -3 & 0 & 5 \\ 6 & -1 & 2 & 4 \end{pmatrix}$

Exercise 1.11.2. Suppose that $\boldsymbol{A} \in \mathbb{R}^{n\times n}$.

(a) If $n = 2$, show that $\det(\boldsymbol{A}^{\mathrm{T}}) = \det(\boldsymbol{A})$.
(b) If $n = 3$, show that $\det(\boldsymbol{A}^{\mathrm{T}}) = \det(\boldsymbol{A})$.
(c) Show that $\det(\boldsymbol{A}^{\mathrm{T}}) = \det(\boldsymbol{A})$ for any n.

Exercise 1.11.3. Suppose that $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$. Show that the matrix product $\boldsymbol{A}\boldsymbol{B}$ is invertible if and only if both $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible. (*Hint*: Use Proposition 1.11.8(b))

Exercise 1.11.4. Suppose that $\boldsymbol{D} = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ is a diagonal matrix, e.g.,

$$\mathrm{diag}(\lambda_1, \lambda_2) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad \text{etc.}$$

(a) If $n = 2$, show that $\det(\boldsymbol{D}) = \lambda_1\lambda_2$.
(b) If $n = 3$, show that $\det(\boldsymbol{D}) = \lambda_1\lambda_2\lambda_3$.
(c) Show that for any n,

$$\det(\boldsymbol{D}) = \prod_{j=1}^{n} \lambda_j.$$

Exercise 1.11.5. Here we generalize the result of Proposition 1.11.8(a). For a matrix $\boldsymbol{A} = (\boldsymbol{a}_1\, \boldsymbol{a}_2\, \boldsymbol{a}_3 \cdots \boldsymbol{a}_n) \in \mathbb{R}^{n\times n}$, let $\boldsymbol{B}$ be defined as $\boldsymbol{B} = (c_1\boldsymbol{a}_1\, c_2\boldsymbol{a}_2\, c_3\boldsymbol{a}_3 \cdots c_n\boldsymbol{a}_n)$.

(a) Show that $\boldsymbol{B} = \boldsymbol{AC}$, where $\boldsymbol{C} = \operatorname{diag}(c_1, c_2, \dots, c_n)$ is a diagonal matrix.
(b) If $n = 2$, show that $\det(\boldsymbol{B}) = c_1 c_2 \det(\boldsymbol{A})$.
(c) Show that for $n \geq 3$,

$$\det(\boldsymbol{B}) = \left(\prod_{j=1}^{n} c_j \right) \det(\boldsymbol{A}).$$

(*Hint*: Use Proposition 1.11.8(b) and Exercise 1.11.4)

Exercise 1.11.6. Suppose that $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is an *upper triangular matrix*, i.e., all of the entries below the diagonal are zero.

(a) Show that $\det(\boldsymbol{A})$ is the product of the diagonal entries. *Hint*: Show that it is true for $n = 2, 3$, and then use an induction argument.
(b) Show that $\boldsymbol{A}^{\mathrm{T}}$ is a *lower triangular matrix*, i.e., all of the entries above the diagonal are zero.
(c) Show that $\det(\boldsymbol{A}^{\mathrm{T}})$ is the product of the diagonal entries. *Hint*: Use the result of Exercise 1.11.2.

Exercise 1.11.7. A matrix $\boldsymbol{V} \in \mathbb{R}^{3\times 3}$ is said to be a Vandermonde matrix if

$$\boldsymbol{V} = \begin{pmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{pmatrix}.$$

(a) Show that $\det(\boldsymbol{V}) = (b - a)(c - a)(c - b)$.
(b) What conditions must the scalars a, b, c satisfy in order that $\boldsymbol{V}$ be invertible?

Exercise 1.11.8. Suppose that

$$\boldsymbol{A}(\lambda) = \begin{pmatrix} 3-\lambda & -2 \\ -2 & 3-\lambda \end{pmatrix}.$$

For which value(s) of λ does the system $\boldsymbol{A}(\lambda)\boldsymbol{x} = \boldsymbol{0}$ have a nontrivial solution? For one such value of λ, compute a corresponding nontrivial solution.

Exercise 1.11.9. Suppose that $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is such that $\boldsymbol{Ax} = \boldsymbol{0}$ has infinitely many solutions. What can be said about $\det(\boldsymbol{A})$? Explain.

Exercise 1.11.10. Determine if each of the following statements is true or false. Provide an explanation for your answer.

(a) If $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ has a pivot in every row, then $\det(\boldsymbol{A}) = 0$.
(b) If $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution for any $\boldsymbol{b}$, then $\det(\boldsymbol{A}) = 0$.
(c) If $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is a diagonal matrix, then $\det(\boldsymbol{A})$ is the product of the diagonal entries.
(d) If the RREF of $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ has one zero row, then $\det(\boldsymbol{A}) \neq 0$.

1.12 Linear algebra with complex-valued numbers, vectors, and matrices

Before we proceed to our last topic on matrices, we will need to understand the basics associated with complex-valued numbers, and the associated algebraic manipulations. As we will see, these will be naturally encountered in future calculations with square matrices, even if the matrix in question contains only real-valued entries.

We say that $z \in \mathbb{C}$ if $z = a + \mathrm{i}b$, where $a, b \in \mathbb{R}$, and $\mathrm{i}^2 = -1$. The number a is the *real part* of the complex number, and is sometimes denoted by $\mathrm{Re}(z)$, i.e., $\mathrm{Re}(z) = a$. The number b is the *imaginary part* of the complex number, and is sometimes denoted by $\mathrm{Im}(z)$, i.e., $\mathrm{Im}(z) = b$. We say a vector $\boldsymbol{v} \in \mathbb{C}^n$ if each entry is complex-valued, and we will often write

$$\boldsymbol{v} = \boldsymbol{p} + \mathrm{i}\boldsymbol{q}, \quad \boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^n.$$

The vector $\boldsymbol{p}$ is the real part of $\boldsymbol{v}$, i.e., $\mathrm{Re}(\boldsymbol{v}) = \boldsymbol{p}$, and the vector $\boldsymbol{q}$ is the imaginary part, i.e., $\mathrm{Im}(\boldsymbol{v}) = \boldsymbol{q}$. For example,

$$\begin{pmatrix} 1 - \mathrm{i}5 \\ 2 + \mathrm{i}7 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mathrm{i}\begin{pmatrix} -5 \\ 7 \end{pmatrix} \quad \Rightarrow \quad \boldsymbol{p} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \ \boldsymbol{q} = \begin{pmatrix} -5 \\ 7 \end{pmatrix}.$$

We say that a matrix $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ if each entry of the matrix is (possibly) complex-valued.

The addition/subtraction of two complex numbers is as expected: add/subtract the real parts and imaginary parts. For example,

$$(2 - \mathrm{i}3) + (3 + \mathrm{i}2) = (2 + 3) + \mathrm{i}(-3 + 2) = 5 - \mathrm{i}.$$

As for multiplication, we multiply products of sums in the usual way, and use the fact that $\mathrm{i}^2 = -1$; for example,

$$\begin{aligned}(2 - \mathrm{i}3)(3 + \mathrm{i}2) &= (2)(3) + (-\mathrm{i}3)(3) + (2)(\mathrm{i}2) + (-\mathrm{i}3)(\mathrm{i}2) \\ &= 6 - \mathrm{i}9 + \mathrm{i}4 - \mathrm{i}^2 6 = 12 - \mathrm{i}5.\end{aligned}$$

In particular, note that

$$c(a + \mathrm{i}b) = ac + \mathrm{i}bc,$$

i.e., multiplication of a complex number by a real number gives a complex number in which the real and imaginary parts of the original number are both multiplied by the real number. For example,

$$7(-4 + \mathrm{i}9) = -28 + \mathrm{i}63.$$

Before we can consider the problem of division, we must first think about the size of a complex number. The *complex-conjugate* of a complex number z, which is denoted by $\overline{z}$, is given by taking the negative of the imaginary part, i.e.,

$$z = a + \mathrm{i}b \quad \Rightarrow \quad \overline{z} = a - \mathrm{i}b.$$

If the number is real-valued, then $\overline{z} = z$. The complex-conjugate of a a vector $\boldsymbol{v} \in \mathbb{C}^n$ is given by

$$\overline{\boldsymbol{v}} = \overline{\boldsymbol{p} + \mathrm{i}\boldsymbol{q}} = \boldsymbol{p} - \mathrm{i}\boldsymbol{q}.$$

The complex-conjugate of a matrix $\boldsymbol{A}$ is written as $\overline{\boldsymbol{A}}$, and the definition is what is to be expected. If $\boldsymbol{A} = (a_{jk})$, then $\overline{\boldsymbol{A}} = (\overline{a_{jk}})$. For example,

$$\boldsymbol{A} = \begin{pmatrix} 2 - \mathrm{i}5 & 3 \\ 1 + \mathrm{i}7 & -3 + \mathrm{i}5 \end{pmatrix} \quad \Rightarrow \quad \overline{\boldsymbol{A}} = \begin{pmatrix} 2 + \mathrm{i}5 & 3 \\ 1 - \mathrm{i}7 & -3 - \mathrm{i}5 \end{pmatrix}.$$

Regarding the conjugate, it is not difficult to check that

$$\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2},$$

i.e., the conjugate of a product is the product of the conjugates (see Exercise 1.12.1(a)). We further have

$$z\overline{z} = (a + \mathrm{i}b)(a - \mathrm{i}b) = a^2 + b^2 > 0,$$

and using this fact we say that the *magnitude* (*absolute value*) of a complex number is

$$|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}.$$

It is not difficult to check that

$$|z_1 z_2| = |z_1||z_2|,$$

i.e., the magnitude of a product is the product of the magnitudes (see Exercise 1.12.1(b)).

We consider the division of two complex numbers by thinking of it as a multiplication problem. We first multiply the complex number by the number one represented as the complex-conjugate of the denominator divided by the complex-conjugate of the denominator. We then write

$$\frac{z_1}{z_2} = \frac{z_1}{z_2}\frac{\overline{z_2}}{\overline{z_2}} = \frac{1}{z_2\overline{z_2}} z_1\overline{z_2} = \frac{1}{|z_2|^2} z_1\overline{z_2},$$

so that division has been replaced by the appropriate multiplication. For example,

$$\frac{2 - \mathrm{i}3}{3 + \mathrm{i}2} = \left(\frac{2 - \mathrm{i}3}{3 + \mathrm{i}2}\right)\left(\frac{3 - \mathrm{i}2}{3 - \mathrm{i}2}\right) = \frac{1}{13}[(2 - \mathrm{i}3)(3 - \mathrm{i}2)] = \frac{1}{13}(-\mathrm{i}13) = -\mathrm{i}.$$

We now derive and state a very important identity - *Euler's formula* - which connects the exponential function to the sine and cosine. This will be accomplished via the use of the Maclaurin series for the exponential and trigonometric functions. Recall that

$$\begin{aligned}
\mathrm{e}^x &= \sum_{j=0}^{\infty} \frac{x^j}{j!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \\
\sin(x) &= \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j+1}}{(2j+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \\
\cos(x) &= \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j}}{(2j)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots,
\end{aligned}$$

and that each series converges for all $x \in \mathbb{R}$. Since

$$\mathrm{i}^2 = -1, \quad \mathrm{i}^3 = \mathrm{i}^2\mathrm{i} = -\mathrm{i}, \quad \mathrm{i}^4 = \mathrm{i}^2\mathrm{i}^2 = 1,$$

we can write for $\theta \in \mathbb{R}$,

$$\begin{aligned}
\mathrm{e}^{\mathrm{i}\theta} &= \sum_{j=0}^{\infty} \frac{(\mathrm{i}\theta)^j}{j!} = 1 + \mathrm{i}\theta - \frac{\theta^2}{2!} - \mathrm{i}\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + \mathrm{i}\frac{\theta^5}{5!} + \cdots \\
&= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + \cdots + (-1)^j \frac{\theta^{2j}}{(2j)!} + \cdots\right) + \\
&\qquad \mathrm{i}\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \cdots + (-1)^j \frac{\theta^{2j+1}}{(2j+1)!} + \cdots\right).
\end{aligned}$$

Noting that the real part is the Maclaurin series for $\cos(\theta)$, and the imaginary part is the Maclaurin series for $\sin(\theta)$, we arrive at Euler's formula,

$$e^{i\theta} = \cos(\theta) + i\sin(\theta).$$

Note that for any $\theta \in \mathbb{R}$,

$$|e^{i\theta}| = \sqrt{\cos^2(\theta) + \sin^2(\theta)} = 1.$$

Further note that Euler's formula yields the intriguing identity,

$$e^{i\pi} = -1,$$

which brings into one simple formula some of the most important constants and concepts in all of mathematics.

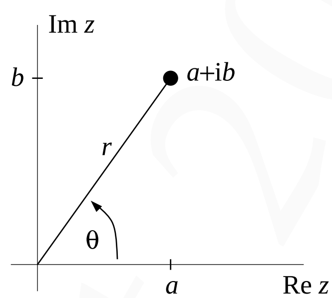


Fig. 1.3 (color online) A cartoon illustration of the polar representation of complex numbers.

As a consequence of Euler's formula we are able to write complex numbers using a polar representation. Let $z = a + ib$ be given. We know that if we represent the point (a,b) in the xy-plane, then the distance from the origin is $r = \sqrt{a^2 + b^2}$, and the angle from the positive x-axis satisfies $\tan(\theta) = b/a$ (see Figure 1.3). This allows us the polar coordinate representation,

$$a = r\cos(\theta), \quad b = r\sin(\theta).$$

Now, we know the magnitude of the complex number is $|z| = \sqrt{a^2 + b^2}$, so we could write

$$a = |z|\cos(\theta), \quad b = |z|\sin(\theta).$$

Upon using Euler's formula we finally see

$$z = a + ib = |z|\cos(\theta) + i|z|\sin(\theta) = |z|[\cos(\theta) + i\sin(\theta)] = |z|e^{i\theta}, \tag{1.12.1}$$

where again

$$|z| = \sqrt{a^2 + b^2}, \quad \tan(\theta) = \frac{b}{a}.$$

As we will see in the case study of Chapter 1.14.3, this representation of a complex number allows us to more easily understand the multiplication of complex-valued numbers.

Does anything really change if we consider the previous linear algebra calculations and concepts under the assumption that the matrices and vectors have complex-valued entries? In summary, no. The definitions and properties of the span of a set of vectors, and subspaces - in particular, the subspaces Null($\boldsymbol{A}$) and Col($\boldsymbol{A}$) - remain the same; indeed, the only difference is that the constants may now be complex-valued. A basis of a subspace is still computed in the same manner, and the dimension of a subspace is still the number of basis vectors. Again, the only difference is that the vectors may have complex-valued entries. As for the inner-product, if we define it for vectors on $\mathbb{C}^n$ as

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{j=1}^{n} x_j \overline{y_j} = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n},$$

then the desired properties will still hold. In particular, we will still have $\langle \boldsymbol{x}, \boldsymbol{x} \rangle$ is a nonnegative real number, and will be zero if and only if $\boldsymbol{x} = \boldsymbol{0}$. Finally, nothing changes for matrix/vector and matrix/matrix multiplication, the calculation of the inverse of a square matrix, and the calculation of the determinant for a square matrix. In conclusion, the only reason we did not start the chapter with a discussion of linear systems with complex-valued coefficients is for the sake of pedagogy, as it is easier to visualize vectors in $\mathbb{R}^n$, and subspaces which are realized as real-valued linear combinations of vectors in $\mathbb{R}^n$.

◁ *Example 1.12.1.* Let us see how we can use our understanding of the algebra of complex numbers when doing Gaussian elimination. Consider the linear system

$$\begin{aligned} (1-\mathrm{i})x_1 + 4x_2 &= 6 \\ (-2+\mathrm{i}3)x_1 + (-8+\mathrm{i}3)x_2 &= -9. \end{aligned}$$

Performing Gaussian elimination on the augmented matrix yields

$$\left(\begin{array}{cc|c} 1-\mathrm{i} & 4 & 6 \\ -2+\mathrm{i}3 & -8+\mathrm{i}3 & -9 \end{array}\right) \xrightarrow{(1/(1-\mathrm{i}))\boldsymbol{\rho}_1} \left(\begin{array}{cc|c} 1 & 2+\mathrm{i}2 & 3+\mathrm{i}3 \\ -2+\mathrm{i}3 & -8+\mathrm{i}3 & -9 \end{array}\right)$$
$$\xrightarrow{(2-\mathrm{i}3)\boldsymbol{\rho}_1+\boldsymbol{\rho}_2} \left(\begin{array}{cc|c} 1 & 2+\mathrm{i}2 & 3+\mathrm{i}3 \\ 0 & -2+\mathrm{i} & 6-\mathrm{i}3 \end{array}\right) \xrightarrow{(1/(-2+\mathrm{i}))\boldsymbol{\rho}_2} \left(\begin{array}{cc|c} 1 & 2+\mathrm{i}2 & 3+\mathrm{i}3 \\ 0 & 1 & -3 \end{array}\right)$$
$$\xrightarrow{(-2-\mathrm{i}2)\boldsymbol{\rho}_2+\boldsymbol{\rho}_1} \left(\begin{array}{cc|c} 1 & 0 & 9+\mathrm{i}9 \\ 0 & 1 & -3 \end{array}\right).$$

The solution is the last column,

$$\boldsymbol{x} = \begin{pmatrix} 9+\mathrm{i}9 \\ -3 \end{pmatrix} = \boldsymbol{x} = \begin{pmatrix} 9 \\ -3 \end{pmatrix} + \mathrm{i}\boldsymbol{x} = \begin{pmatrix} 9 \\ 0 \end{pmatrix}.$$

◁ *Example 1.12.2.* For another example, let us find Null($\boldsymbol{A}$) for the matrix

$$\boldsymbol{A} = \begin{pmatrix} 3-\mathrm{i} & 4 \\ 5 & 6+\mathrm{i}2 \end{pmatrix}.$$

Since

$$A \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & (6+\mathrm{i}2)/5 \\ 0 & 0 \end{pmatrix},$$

the null space is found by solving

$$x_1 + \frac{6+\mathrm{i}2}{5}x_2 = 0.$$

Upon setting $x_2 = 5t$ the solution vector is given by

$$\boldsymbol{x} = \begin{pmatrix} -(6+\mathrm{i}2)t \\ 5t \end{pmatrix} = t\begin{pmatrix} -6-\mathrm{i}2 \\ 5 \end{pmatrix}.$$

We conclude that

$$\mathrm{Null}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} -6-\mathrm{i}2 \\ 5 \end{pmatrix}\}, \quad \dim[\mathrm{Null}(\boldsymbol{A})] = 1.$$

◁ *Example 1.12.3.* Let us find those vectors $\boldsymbol{b}$ for which the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent with

$$\boldsymbol{A} = \begin{pmatrix} 2-\mathrm{i} & 4 \\ 5 & 8+\mathrm{i}4 \end{pmatrix}.$$

Gaussian elimination yields that the RREF of $\boldsymbol{A}$ is

$$\boldsymbol{A} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & (8+\mathrm{i}4)/5 \\ 0 & 0 \end{pmatrix}.$$

Since the RREF of $\boldsymbol{A}$ has a zero row, the system $\boldsymbol{Ax} = \boldsymbol{b}$ is not consistent for all $\boldsymbol{b}$. Moreover, only the first column is a pivot column, so by using Lemma 1.7.8 we know that a basis for $\mathrm{Col}(\boldsymbol{A})$ is the first column of $\boldsymbol{A}$, i.e.,

$$\mathrm{Col}(\boldsymbol{A}) = \mathrm{Span}\{\begin{pmatrix} 2-\mathrm{i} \\ 5 \end{pmatrix}\}, \quad \mathrm{rank}(\boldsymbol{A}) = 1.$$

The linear system is consistent if and only if $\boldsymbol{b} \in \mathrm{Col}(\boldsymbol{A})$.

Exercises

Exercise 1.12.1. Let $z_1 = a_1 + \mathrm{i}b_1$ and $z_2 = a_2 + \mathrm{i}b_2$ be two complex numbers. Show that

(a) $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$
(b) $|z_1 z_2| = |z_1||z_2|$.

Exercise 1.12.2. Write each complex number z in the form $|z|\mathrm{e}^{\mathrm{i}\theta}$, where $-\pi < \theta \le \pi$.

(a) $z = 3 - \mathrm{i}4$
(b) $z = -2 + \mathrm{i}5$
(c) $z = -3 - \mathrm{i}7$
(d) $z = 6 + \mathrm{i}$

Exercise 1.12.3. Solve each system of equations, or explain why no solution exists.

(a) $(3-\mathrm{i})x_1 + 2x_2 = 2,\ -4x_1 + (1+\mathrm{i}4)x_2 = -3$
(b) $x_1 + (-2+\mathrm{i}5)x_2 = -3,\ (1-\mathrm{i}5)x_1 + 3x_2 = 12$

Exercise 1.12.4. For each of the below problems compute the product $\boldsymbol{A}\boldsymbol{x}$ when it is well-defined. If the product cannot be computed, explain why.

(a) $\boldsymbol{A} = \begin{pmatrix} 2+\mathrm{i} & 3 \\ -2 & 1+\mathrm{i}4 \\ 3 & 7 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} 2+\mathrm{i}3 \\ 8 \end{pmatrix}$

(b) $\boldsymbol{A} = \begin{pmatrix} 2 & -1+\mathrm{i}3 & -4 \\ 2+\mathrm{i}5 & 6 & 3-\mathrm{i}7 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} 2 \\ 9 \\ 4+\mathrm{i}3 \end{pmatrix}.$

Exercise 1.12.5. For each matrix $\boldsymbol{A}$, find $\mathrm{Null}(\boldsymbol{A})$, and determine its dimension.

(a) $\boldsymbol{A} = \begin{pmatrix} 2+\mathrm{i}3 & 26 \\ 2 & 8-\mathrm{i}12 \end{pmatrix}$

(b) $\boldsymbol{A} = \begin{pmatrix} 1-\mathrm{i}4 & 17 \\ 2 & 2+\mathrm{i}8 \end{pmatrix}$

Exercise 1.12.6. Solve the following linear system, and explicitly identify the homogeneous solution, $\boldsymbol{x}_{\mathrm{h}}$, and the particular solution, $\boldsymbol{x}_{\mathrm{p}}$:

$$\begin{pmatrix} 3+\mathrm{i}2 & -26 \\ -2 & 12-\mathrm{i}8 \end{pmatrix}\boldsymbol{x} = \begin{pmatrix} 13 \\ -6+\mathrm{i}4 \end{pmatrix}.$$

Exercise 1.12.7. In Example 1.7.6 it was shown that $\dim[\mathbb{R}^n] = n$. Show that $\dim[\mathbb{C}^n] = n$.

1.13 Eigenvalues and eigenvectors

Consider a square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. As we will see in the two case studies in Chapter 1.14, as well as when solving homogeneous systems of ODEs in Chapter 3, it will be especially useful to identify a set of vectors, say $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$, such that for each vector there is a constant λ_j such that

$$\boldsymbol{A}\boldsymbol{v}_j = \lambda_j \boldsymbol{v}_j, \quad j = 1, \ldots, n. \tag{1.13.1}$$

The vectors $\boldsymbol{v}_j$, which are known as *eigenvectors*, and multiplicative factors λ_j, which are known as *eigenvalues*, may be complex-valued (see Chapter 1.12). If the eigenvalues are complex-valued, then the corresponding eigenvector also has complex-valued entries. Eigenvectors are vectors that have the property that matrix multiplication by $\boldsymbol{A}$ leads to a scalar multiple of the original vector.

1.13.1 Characterization of eigenvalues and eigenvectors

How do we find these vectors $\boldsymbol{v}$ and associated multiplicative factors λ? We can rewrite 1.13.1 as

$$\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v} \quad (\boldsymbol{A} - \lambda\boldsymbol{I}_n)\boldsymbol{v} = \boldsymbol{0}.$$

Recalling the definition of a null space, an eigenvector $\boldsymbol{v}$ can be found if we can find an eigenvalue λ such that

$$\dim[\mathrm{Null}(\boldsymbol{A} - \lambda\boldsymbol{I}_n)] \geq 1.$$

If λ is an eigenvalue, then we will call $\mathrm{Null}(\boldsymbol{A}-\lambda\boldsymbol{I}_n)$ the *eigenspace*. An eigenvector is any basis vector of the eigenspace. Since a basis is not unique, neither will be the eigenvectors (recall the discussion in Chapter 1.7). However, the *number* of basis vectors is unique (recall Lemma 1.7.4), so associated with each eigenvalue there will be a fixed number of linearly independent eigenvectors.

Now, if we are given an eigenvalue, then it is straightforward to compute a basis for the associated eigenspace. The problem really is in finding an eigenvalue. This requires an additional equation, for at the moment the linear system is a set of n equations with $n+1$ variables, which are the n components of the vector plus the associated eigenvalue. In constructing this additional equation we can rely upon the result of Theorem 1.11.9, in which it is stated that a square matrix has a nontrivial null space if and only if its determinant is zero. If we set

$$p_{\boldsymbol{A}}(\lambda) = \det(\boldsymbol{A}-\lambda\boldsymbol{I}_n),$$

then the eigenvalues will correspond to the zeros of the *characteristic polynomial* $p_{\boldsymbol{A}}(\lambda)$. While we will not do it here, it is not difficult to show that the characteristic polynomial is a polynomial of degree n, the size of the square matrix (see Exercise 1.13.4).

We summarize this discussion with the following result:

Theorem 1.13.1. *Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. The zeros of the n^{th}-order characteristic polynomial $p_{\boldsymbol{A}}(\lambda) = \det(\boldsymbol{A}-\lambda\boldsymbol{I}_n)$ are the eigenvalues of the matrix $\boldsymbol{A}$. The (not unique) eigenvectors associated with an eigenvalue are a basis for $\mathrm{Null}(\boldsymbol{A}-\lambda\boldsymbol{I}_n)$.*

Before going any further in the discussion of the theory associated with eigenvalues and eigenvectors, let us do a relatively simple computation.

◁ *Example 1.13.2.* Let us find the eigenvalues and associated eigenvectors for

$$\boldsymbol{A} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}.$$

We have

$$\boldsymbol{A}-\lambda\boldsymbol{I}_2 = \begin{pmatrix} 3-\lambda & 2 \\ 2 & 3-\lambda \end{pmatrix},$$

so that the characteristic polynomial is

$$p_{\boldsymbol{A}}(\lambda) = (3-\lambda)^2 - 4.$$

The zeros of the characteristic polynomial are $\lambda = 1, 5$. As for the associated eigenvectors, we must compute a basis for $\mathrm{Null}(\boldsymbol{A}-\lambda\boldsymbol{I}_2)$ for each eigenvalue. For the eigenvalue $\lambda_1 = 1$ we have

$$\boldsymbol{A}-\boldsymbol{I}_2 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},$$

which corresponds to the linear equation $v_1 + v_2 = 0$. Since

$$\mathrm{Null}(\boldsymbol{A}-\boldsymbol{I}_2) = \mathrm{Span}\{\begin{pmatrix} -1 \\ 1 \end{pmatrix}\},$$

we have that an associated eigenvector is

$$\lambda_1 = 1; \quad \boldsymbol{v}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Because eigenvectors are not unique, any multiple of $\boldsymbol{v}_1$ given above would be an eigenvector associated with the eigenvalue $\lambda_1 = 1$. For the eigenvalue $\lambda_2 = 5$ we have

$$\boldsymbol{A} - 5\boldsymbol{I}_2 \stackrel{\text{RREF}}{\longrightarrow} \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix},$$

which corresponds to the linear equation $v_1 - v_2 = 0$. Since

$$\text{Null}(\boldsymbol{A} - 5\boldsymbol{I}_2) = \text{Span}\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\},$$

we have that an associated eigenvector is

$$\lambda_2 = 5; \quad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Before continuing, we need to decide how many eigenvectors are to be associated with a given eigenvalue. An eigenvalue λ_0 is such that

$$\text{m}_\text{g}(\lambda_0) := \dim[\text{Null}(\boldsymbol{A} - \lambda_0 \boldsymbol{I}_n)] \geq 1.$$

The integer $\text{m}_\text{g}(\lambda_0)$ is the *geometric multiplicity* of the eigenvalue λ_0. We know from Chapter 1.7 that $\text{m}_\text{g}(\lambda_0)$ will be the number of free variables associated with the RREF of $\boldsymbol{A} - \lambda_0 \boldsymbol{I}_n$. Consequently, any basis of the the eigenspace will have $\text{m}_\text{g}(\lambda_0)$ vectors. Since a basis is not unique, the eigenvectors will not be unique; however, once a set of eigenvectors has been chosen, any other eigenvector must be a linear combination of the chosen set.

Since the characteristic polynomial $p_{\boldsymbol{A}}(\lambda)$ is an n^{th}-order polynomial, by the Fundamental Theorem of Algebra it can be factored as

$$p_{\boldsymbol{A}}(\lambda) = c(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n), \quad c \neq 0.$$

If $\lambda_i \neq \lambda_j$ for all $i \neq j$, then all of the eigenvalues are said to be *algebraically simple* (simple). If an eigenvalue λ_j is algebraically simple, then the associated eigenspace will be one-dimensional, i.e., $\text{m}_\text{g}(\lambda_j) = 1$, so that all basis vectors for the eigenspace will be scalar multiples of each other. In other words, for simple eigenvalues an associated eigenvector will be a scalar multiple of any other associated eigenvector. If an eigenvalue is not simple, then we will call it a *multiple eigenvalue*. For example, if

$$p_{\boldsymbol{A}}(\lambda) = (\lambda + 1)(\lambda - 1)^2(\lambda - 3)^4,$$

then $\lambda = -1$ is a simple eigenvalue, and $\lambda = 1$ and $\lambda = 3$ are multiple eigenvalues. The *(algebraic) multiplicity* of a multiple eigenvalue is the order of the zero of the characteristic polynomial, and will be denoted by $\text{m}_\text{a}(\lambda_0)$. In this example $\lambda = -1$ is such that $\text{m}_\text{a}(-1) = 2$ (a double eigenvalue), and $\lambda = 3$ is such that $\text{m}_\text{a}(3) = 4$ (a quartic eigenvalue).

It is a fundamental fact of linear algebra that the two multiplicities are related via

$$1 \leq \text{m}_\text{g}(\lambda_0) \leq \text{m}_\text{a}(\lambda_0).$$

As already stated, if an eigenvalue is algebraically simple, i.e., $\mathrm{m_a}(\lambda_0) = 1$, then it must be true that $\mathrm{m_g}(\lambda_0) = \mathrm{m_a}(\lambda_0) = 1$. On the other hand, if $\mathrm{m_a}(\lambda_0) \geq 2$ it may be the case that $\mathrm{m_g}(\lambda_0) < \mathrm{m_a}(\lambda_0)$. This situation is nongeneric, and can be rectified by a small perturbation of the matrix $\boldsymbol{A}$. Indeed, it will generically be the case that all of the eigenvalues for a given matrix are algebraically simple.

As a final remark, we remind the reader that while eigenvectors themselves are not unique, the number of eigenvectors is unique. In all that follows we will compute only a set of eigenvectors associated with a particular eigenvalue, and not spend much effort discussing the associated eigenspace. The reader needs to always remember that for a given set of eigenvectors any linear combination of the given eigenvectors also counts as an eigenvector.

◁ *Example 1.13.3.* Let us find the eigenvalues and associated eigenvectors for

$$\boldsymbol{A} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 2 & 3 \end{pmatrix}.$$

We have

$$\boldsymbol{A} - \lambda \boldsymbol{I}_3 = \begin{pmatrix} 5-\lambda & 0 & 0 \\ 0 & 3-\lambda & 2 \\ 0 & 2 & 3-\lambda \end{pmatrix},$$

so that the characteristic polynomial is

$$p_{\boldsymbol{A}}(\lambda) = (5-\lambda)[(3-\lambda)^2 - 4].$$

The zeros of the characteristic polynomial are $\lambda = 1, 5$, where $\lambda = 5$ is a double root, i.e., $\lambda = 5$ is a double eigenvalue. Regarding the algebraic multiplicities we have $\mathrm{m_a}(1) = 1$ and $\mathrm{m_a}(5) = 2$. As for the associated eigenvectors, we have for the eigenvalue $\lambda_1 = 1$

$$\boldsymbol{A} - 1\boldsymbol{I}_3 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

which corresponds to the linear system $v_1 = 0$, $v_2 + v_3 = 0$. An eigenvector is then given by

$$\lambda_1 = 1; \quad \boldsymbol{v}_1 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

For the eigenvalue $\lambda_2 = \lambda_3 = 5$ we have

$$\boldsymbol{A} - 5\boldsymbol{I}_3 \xrightarrow{\text{RREF}} \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

which corresponds to the linear equation $v_2 - v_3 = 0$. There are two free variables, v_1 and v_3, so that there are two eigenvectors, which are given by

$$\lambda_2 = \lambda_3 = 5; \quad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

◁ *Example* 1.13.4. Let us find the eigenvalues and associated eigenvectors for

$$\boldsymbol{A} = \begin{pmatrix} 5 & 0 & 8 \\ 0 & 3 & 2 \\ 0 & 2 & 3 \end{pmatrix}.$$

Note that only one entry in $\boldsymbol{A}$, a_{13}, has changed from the previous example. We have

$$\boldsymbol{A} - \lambda \boldsymbol{I}_3 = \begin{pmatrix} 5-\lambda & 0 & 8 \\ 0 & 3-\lambda & 2 \\ 0 & 2 & 3-\lambda \end{pmatrix},$$

so that the characteristic polynomial is again

$$p_{\boldsymbol{A}}(\lambda) = (5-\lambda)[(3-\lambda)^2 - 4].$$

As in the previous example, the eigenvalues are $\lambda = 1$ and $\lambda = 5$ with $\mathrm{m_a}(1) = 1$ and $\mathrm{m_a}(5) = 2$. As for the associated eigenvectors, we have for the eigenvalue $\lambda_1 = 1$

$$\boldsymbol{A} - 1\boldsymbol{I}_3 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

which corresponds to the linear system $v_1 + 2v_3 = 0$, $v_2 + v_3 = 0$. The eigenvector is then given by

$$\lambda_1 = 1; \quad \boldsymbol{v}_1 = \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix}.$$

For the eigenvalue $\lambda_2 = \lambda_3 = 5$ we have

$$\boldsymbol{A} - 5\boldsymbol{I}_3 \xrightarrow{\text{RREF}} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

which corresponds to the linear system $v_2 = v_3 = 0$. Unlike the previous example there is now only one free variable, v_1, which means that there is only one eigenvector associated with both of these eigenvalues,

$$\lambda_2 = \lambda_3 = 5; \quad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Thus, in this example there are not as many eigenvectors as there are eigenvalues.

◁ *Example* 1.13.5. Let us find the eigenvalues and associated eigenvectors for

$$\boldsymbol{A} = \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}.$$

We have

$$\boldsymbol{A} - \lambda \boldsymbol{I}_2 = \begin{pmatrix} 3-\lambda & -2 \\ 2 & 3-\lambda \end{pmatrix},$$

so that the characteristic polynomial is

$$p_A(\lambda) = (3-\lambda)^2 + 4.$$

The zeros of the characteristic polynomial are $\lambda = 3 \pm \mathrm{i}2$. Note that this set of eigenvalues is a complex-conjugate pair. As for the associated eigenvectors, we have for the eigenvalue $\lambda_1 = 3 + \mathrm{i}2$

$$\boldsymbol{A} - (3+\mathrm{i}2)\boldsymbol{I}_2 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -\mathrm{i} \\ 0 & 0 \end{pmatrix},$$

which corresponds to the linear equation $v_1 - \mathrm{i}v_2 = 0$. An eigenvector is then given by

$$\lambda_1 = 3 + \mathrm{i}2; \quad \boldsymbol{v}_1 = \begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mathrm{i}\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

For the eigenvalue $\lambda_2 = 3 - \mathrm{i}2$ we have

$$\boldsymbol{A} - (3-\mathrm{i}2)\boldsymbol{I}_2 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & \mathrm{i} \\ 0 & 0 \end{pmatrix},$$

which corresponds to the linear equation $v_1 + \mathrm{i}v_2 = 0$. An eigenvector is then given by

$$\lambda_2 = 3 - \mathrm{i}2; \quad \boldsymbol{v}_2 = \begin{pmatrix} -\mathrm{i} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \mathrm{i}\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

As was the case for the eigenvalues, the eigenvectors also come in a complex-conjugate pair.

◁ *Example 1.13.6.* Let us find the eigenvalues and associated eigenvectors for

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ -5 & -2 \end{pmatrix}.$$

We have

$$\boldsymbol{A} - \lambda\boldsymbol{I}_2 = \begin{pmatrix} -\lambda & 1 \\ -5 & -2-\lambda \end{pmatrix},$$

so that the characteristic polynomial is

$$p_A(\lambda) = \lambda^2 + 2\lambda + 5 = (\lambda+1)^2 + 4.$$

The zeros of the characteristic polynomial are $\lambda = -1 \pm \mathrm{i}2$. Note that once again the eigenvalues arise in a complex-conjugate pair. As for the associated eigenvectors, we have for the eigenvalue $\lambda_1 = -1 + \mathrm{i}2$

$$\boldsymbol{A} - (-1+\mathrm{i}2)\boldsymbol{I}_2 \xrightarrow{\text{RREF}} \begin{pmatrix} 1-\mathrm{i}2 & 1 \\ 0 & 0 \end{pmatrix},$$

which corresponds to the linear equation $(1-\mathrm{i}2)v_1 + v_2 = 0$. An eigenvector is then given by

$$\lambda_1 = -1 + \mathrm{i}2; \quad \boldsymbol{v}_1 = \begin{pmatrix} -1+\mathrm{i}2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \mathrm{i}\begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

For the eigenvalue $\lambda_2 = -1 - \mathrm{i}2$ we eventually see that an eigenvector is given by

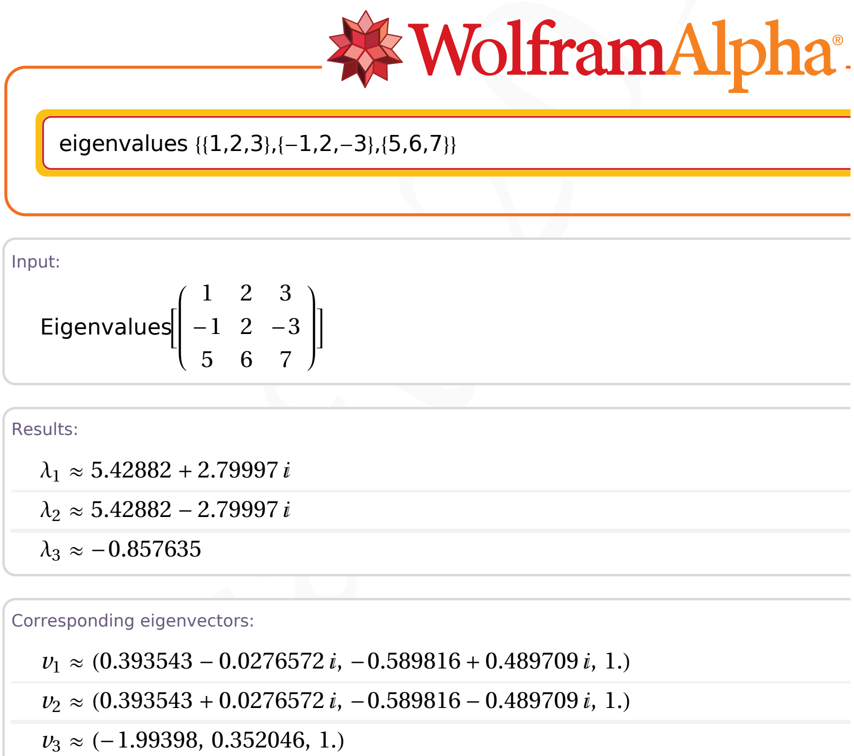
$$\boldsymbol{v}_2 = \begin{pmatrix} -1 - \mathrm{i}2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} - \mathrm{i} \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Thus, just as in the previous example the associated eigenvectors also come in a complex-conjugate pair.

◁ *Example* 1.13.7. We finally consider an example for which the eigenvalues and eigenvectors must be computed numerically. Here

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & -3 \\ 5 & 6 & 7 \end{pmatrix} \in \mathbb{R}^{3\times 3},$$

which means that $p_A(\lambda)$ is a third-order polynomial. Unless the problem is very special, it is generally the case that it is not possible to (easily) find the three roots. Using WolframAlpha we get



In other words,

$$\lambda_1 \sim 5.43 + \mathrm{i}2.80,\ \boldsymbol{v}_2 \sim \begin{pmatrix} 0.39 \\ -0.59 \\ 1.00 \end{pmatrix} + \mathrm{i} \begin{pmatrix} -0.03 \\ 0.49 \\ 0.00 \end{pmatrix};\ \lambda_3 \sim -0.86,\ \boldsymbol{v}_3 \sim \begin{pmatrix} -1.99 \\ 0.35 \\ 1.00 \end{pmatrix}.$$

The second eigenvalue is the complex-conjugate conjugate of the second, i.e., $\lambda_2 = \overline{\lambda_1}$, and the associated eigenvector is the complex-conjugate of $\boldsymbol{v}_1$, i.e., $\boldsymbol{v}_2 = \overline{\boldsymbol{v}_1}$.

1.13.2 Properties

The last three examples highlight a general phenomena. Suppose that $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, and suppose that $\lambda = a + \mathrm{i}b$ is an eigenvalue with associated eigenvector $\boldsymbol{v} = \boldsymbol{p} + \mathrm{i}\boldsymbol{q}$. In other words, suppose that

$$\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v} \quad \Rightarrow \quad \boldsymbol{A}(\boldsymbol{p} + \mathrm{i}\boldsymbol{q}) = (a + \mathrm{i}b)(\boldsymbol{p} + \mathrm{i}\boldsymbol{q}).$$

Taking the complex-conjugate of both sides, and using the fact that the conjugate of a product is the product of the conjugates,

$$\overline{\boldsymbol{A}\boldsymbol{v}} = \overline{\boldsymbol{A}}\,\overline{\boldsymbol{v}}, \quad \overline{\lambda\boldsymbol{v}} = \overline{\lambda}\,\overline{\boldsymbol{v}},$$

gives

$$\overline{\boldsymbol{A}}\,\overline{\boldsymbol{v}} = \overline{\lambda}\,\overline{\boldsymbol{v}}.$$

Since $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\overline{\boldsymbol{A}} = \boldsymbol{A}$, so we conclude

$$\boldsymbol{A}\overline{\boldsymbol{v}} = \overline{\lambda}\overline{\boldsymbol{v}}.$$

This equation is another eigenvalue/eigenvector equation for the matrix $\boldsymbol{A}$. The eigenvalue and associated eigenvector for this equation are related to the original via complex-conjugation. In conclusion, if $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, then the eigenvalues come in the complex-conjugate pairs $\{\lambda, \overline{\lambda}\}$, i.e., $a \pm \mathrm{i}b$, as do the associated eigenvectors $\{\boldsymbol{v}, \overline{\boldsymbol{v}}\}$, i.e., $\boldsymbol{p} \pm \mathrm{i}\boldsymbol{q}$.

We conclude with some additional facts about eigenvalues and eigenvectors of $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, each of which will be useful in applications. First, the eigenvalues tell us something about the invertibility of the matrix. We first observe that by setting $\lambda = 0$,

$$p_{\boldsymbol{A}}(0) = \det(\boldsymbol{A});$$

thus, $\lambda = 0$ is an eigenvalue if and only if $\det(\boldsymbol{A}) = 0$. Since by Theorem 1.11.9 $\boldsymbol{A}$ is invertible if and only if $\det(\boldsymbol{A}) \neq 0$, we have that $\boldsymbol{A}$ is invertible if and only if $\lambda = 0$ is not an eigenvalue. From Exercise 1.13.4(c) we know that the characteristic polynomial is of degree n; hence, by the Fundamental Theorem of Algebra there are precisely n eigenvalues. As we have seen in the previous examples, there may or may not be n linearly independent eigenvectors. However, if the eigenvalues are *distinct* (each one is algebraically simple), then the n eigenvectors are indeed linearly independent. Since $\dim[\mathbb{C}^n] = n$ (see Exercise 1.12.7), this means that we can use the eigenvectors as a basis for $\mathbb{C}^n$.

Theorem 1.13.8. *Consider the matrix* $\boldsymbol{A} \in \mathbb{R}^{n\times n}$.

(a) *If* $\lambda = a + \mathrm{i}b$ *is an eigenvalue with associated eigenvector* $\boldsymbol{v} = \boldsymbol{p} + \mathrm{i}\boldsymbol{q}$ *for some vectors* $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^n$, *then the complex-conjugate* $\overline{\lambda} = a - \mathrm{i}b$ *is an eigenvalue with associated complex-conjugated eigenvector* $\overline{\boldsymbol{v}} = \boldsymbol{p} - \mathrm{i}\boldsymbol{q}$.
(b) $\lambda = 0$ *is an eigenvalue if and only if* $\det(\boldsymbol{A}) = 0$.
(c) $\boldsymbol{A}$ *is invertible if and only if all of the eigenvalues are nonzero.*
(d) *If all the eigenvalues are distinct, i.e., all of the roots of the characteristic polynomial are simple, then the set of eigenvectors forms a basis for* $\mathbb{C}^n$.

If in this proof we remove the assumption that all the eigenvalues are real-valued, the only thing that changes is that we can no longer order the eigenvalues to be increasing

Proof. We only need to show that if the eigenvalues are distinct, then the set of eigenvectors forms a basis. For the sake of exposition, and without loss of generality, we will assume that all of the eigenvalues are real-valued, and so we will order them as $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. The corresponding eigenvectors will be denoted by $\boldsymbol{v}_j$, so that $\boldsymbol{A}\boldsymbol{v}_j = \lambda_j \boldsymbol{v}_j$ for $j = 1, \ldots, n$.

We start by assuming that the eigenvectors are linearly dependent; thus, there will be a nonzero vector $\boldsymbol{c} \in \mathbb{R}^n$ such that

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{0}.$$

Multiplying both sides by $\boldsymbol{A}$ and using the linearity of matrix/vector multiplication gives

$$c_1\boldsymbol{A}\boldsymbol{v}_1 + c_2\boldsymbol{A}\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{A}\boldsymbol{v}_n = \boldsymbol{0} \quad \Rightarrow \quad c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \cdots + c_n\lambda_n\boldsymbol{v}_n = \boldsymbol{0}.$$

Multiplying the first equation by λ_n and subtracting this from the second equation yields

$$c_1(\lambda_n - \lambda_1)\boldsymbol{v}_1 + c_2(\lambda_n - \lambda_2)\boldsymbol{v}_2 + \cdots + c_{n-1}(\lambda_n - \lambda_{n-1})\boldsymbol{v}_{n-1} = \boldsymbol{0}.$$

Since the eigenvalues are distinct, i.e., $\lambda_n - \lambda_j > 0$ for $j = 1, \ldots, n-1$, we now see that the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{n-1}$ are linearly dependent. In conclusion, the set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ being linearly dependent implies that the subset $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{n-1}\}$ is also linearly dependent.

Continuing in this fashion, we eventually conclude that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are linearly dependent, so that there is a nonzero constant c such that $\boldsymbol{v}_2 = c\boldsymbol{v}_1$. Thus,

$$c\lambda_2\boldsymbol{v}_1 = \lambda_2\boldsymbol{v}_2 = \boldsymbol{A}\boldsymbol{v}_2 = c\boldsymbol{A}\boldsymbol{v}_1 = c\lambda_1\boldsymbol{v}_1 \quad \Rightarrow \quad c(\lambda_2 - \lambda_1)\boldsymbol{v}_1 = \boldsymbol{0}.$$

Since $c \neq 0$ and $\lambda_2 > \lambda_1$, this implies $\boldsymbol{v}_1 = \boldsymbol{0}$. But, this statement cannot be true, as eigenvectors must be nonzero. In conclusion, we have contradicted our initial assumption that the eigenvectors are linearly dependent, which means that they are linearly independent. □

1.13.3 Eigenvectors as a basis, and Fourier expansions

The moniker Fourier appeared in Chapter 1.8 when we discussed orthonormal bases. We use it again here because in some instances - e.g., $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$ - the eigenvectors can be scaled to form an orthonormal basis.

As we will see in the upcoming case studies, as well as our study of linear systems of ODEs, it is extremely beneficial to write a given vector in terms of the eigenvectors of a given matrix. For a given matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, suppose that a set of eigenvectors, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$, forms a basis. We know from Theorem 1.13.8(d) that this is possible if the eigenvalues are distinct. Going back to our discussion in Chapter 1.7 we then know that any vector $\boldsymbol{x} \in \mathbb{C}^n$ can be uniquely written through the expansion

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n. \tag{1.13.2}$$

Such an expansion in terms of the eigenvectors is sometimes called a Fourier expansion, and the weights are sometimes called the Fourier coefficients. The Fourier coefficients are found through the solution of the linear system, i.e.,

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{P}\boldsymbol{c}, \quad \boldsymbol{P} = (\boldsymbol{v}_1\ \boldsymbol{v}_2\ \cdots\ \boldsymbol{v}_n),$$

means

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{c} \quad \Rightarrow \quad \boldsymbol{c} = \boldsymbol{P}^{-1}\boldsymbol{x}.$$

The matrix $\boldsymbol{P}$ is invertible because the eigenvectors are assumed to be linearly independent (otherwise, they would not form a basis), and the result of Theorem 1.11.10 states that the inverse exists if and only if the matrix has full rank.

By writing a given vector through a Fourier expansion we develop greater insight into the geometry associated with matrix/vector multiplication. Multiplying both sides of (1.13.2) by $\boldsymbol{A}$ and using linearity gives

$$\boldsymbol{A}\boldsymbol{x} = c_1\boldsymbol{A}\boldsymbol{v}_1 + c_2\boldsymbol{A}\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{A}\boldsymbol{v}_n.$$

Since each vector $\boldsymbol{v}_j$ is an eigenvector with associated eigenvalue λ_j, i.e., $\boldsymbol{A}\boldsymbol{v}_j = \lambda_j\boldsymbol{v}_j$, we can rewrite the right-hand side of the above as

$$c_1\boldsymbol{A}\boldsymbol{v}_1 + c_2\boldsymbol{A}\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{A}\boldsymbol{v}_n = c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \cdots + c_n\lambda_n\boldsymbol{v}_n.$$

Putting the pieces together yields

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n \quad \Rightarrow \quad \boldsymbol{A}\boldsymbol{x} = c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \cdots + c_n\lambda_n\boldsymbol{v}_n.$$

Thus, the Fourier coefficients associated with the vector $\boldsymbol{A}\boldsymbol{x}$ are a scaling of those for the vector $\boldsymbol{x}$, where the scaling is the eigenvalue associated with the eigenvector.

Lemma 1.13.9. *Suppose that for $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ there is a set of linearly independent eigenvectors, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ (guaranteed if the eigenvalues are distinct). For any $\boldsymbol{x} \in \mathbb{C}^n$ there is the Fourier expansion*

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n,$$

where the Fourier coefficients $c_1, c_2, \dots, c_n$ are uniquely determined. Moreover, the vector $\boldsymbol{A}\boldsymbol{x}$ has the Fourier expansion

$$\boldsymbol{A}\boldsymbol{x} = c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \cdots + c_n\lambda_n\boldsymbol{v}_n,$$

where each λ_j is the eigenvalue associated with the eigenvector $\boldsymbol{v}_j$.

◁ *Example 1.13.10.* Consider the matrix

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ -8 & -6 \end{pmatrix}.$$

It can be checked that the eigenvalues and associated eigenvectors are

$$\lambda_1 = -2,\ \boldsymbol{v}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}; \quad \lambda_2 = -4,\ \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ -4 \end{pmatrix}.$$

The eigenvectors are clearly linearly independent, so they form a basis. For a particular example, let us find the Fourier coefficients for the vector $\boldsymbol{x} = (2\ -7)^{\mathrm{T}}$. Using (1.13.2) we have

$$\begin{pmatrix} 2 \\ -7 \end{pmatrix} = c_1\begin{pmatrix} 1 \\ -2 \end{pmatrix} + c_2\begin{pmatrix} 1 \\ -4 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & -4 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

The solution to this linear system is

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & -4 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ -7 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} -4 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -7 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

In other words, the Fourier coefficients are

$$c_1 = \frac{1}{2}, \quad c_2 = \frac{3}{2},$$

so

$$\boldsymbol{x} = \frac{1}{2}\boldsymbol{v}_1 + \frac{3}{2}\boldsymbol{v}_2.$$

Using the result of Lemma 1.13.9 we also know that

$$\boldsymbol{A}\boldsymbol{x} = \frac{1}{2}(-2)\boldsymbol{v}_1 + \frac{3}{2}(-4)\boldsymbol{v}_2 = -\boldsymbol{v}_1 - 6\boldsymbol{v}_2.$$

Exercises

Exercise 1.13.1. Suppose that for a given $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ there is a set of linearly independent eigenvectors, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$. Suppose that a given $\boldsymbol{x}$ has the Fourier expansion

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n.$$

Defining

$$\boldsymbol{A}^\ell := \underbrace{\boldsymbol{A} \cdot \boldsymbol{A} \cdots \boldsymbol{A}}_{\ell \text{ times}},$$

show that:

(a) $\boldsymbol{A}^2\boldsymbol{x} = c_1\lambda_1^2\boldsymbol{v}_1 + c_2\lambda_2^2\boldsymbol{v}_2 + \cdots + c_n\lambda_n^2\boldsymbol{v}_n$
(b) $\boldsymbol{A}^3\boldsymbol{x} = c_1\lambda_1^3\boldsymbol{v}_1 + c_2\lambda_2^3\boldsymbol{v}_2 + \cdots + c_n\lambda_n^3\boldsymbol{v}_n$
(c) if $\ell \geq 4$, $\boldsymbol{A}^\ell\boldsymbol{x} = c_1\lambda_1^\ell\boldsymbol{v}_1 + c_2\lambda_2^\ell\boldsymbol{v}_2 + \cdots + c_n\lambda_n^\ell\boldsymbol{v}_n$ (*Hint*: use (a), (b), and induction).

Exercise 1.13.2. Compute by hand the eigenvalues and all corresponding eigenvectors for each matrix. If the eigenvalue is complex-valued, write the eigenvector in the form $\boldsymbol{v} = \boldsymbol{p} + \mathrm{i}\boldsymbol{q}$.

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}$
(b) $\boldsymbol{A} = \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}$
(c) $\boldsymbol{A} = \begin{pmatrix} 2 & -6 \\ 3 & -7 \end{pmatrix}$
(d) $\boldsymbol{A} = \begin{pmatrix} 1 & 4 \\ -2 & -3 \end{pmatrix}$
(e) $\boldsymbol{A} = \begin{pmatrix} 2 & 5 & -2 \\ 5 & 2 & 1 \\ 0 & 0 & -3 \end{pmatrix}$
(f) $\boldsymbol{A} = \begin{pmatrix} 3 & 0 & 0 \\ 1 & -1 & 2 \\ -2 & -2 & -1 \end{pmatrix}$

Exercise 1.13.3. In each of the following the characteristic polynomial of $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is given. Determine n, list each eigenvalue and its algebraic multiplicity, and state whether or not the matrix is invertible.

(a) $p_{\boldsymbol{A}}(\lambda) = (\lambda-3)(\lambda^2+2\lambda+5)(\lambda-4)^2$
(b) $p_{\boldsymbol{A}}(\lambda) = \lambda^2(\lambda+3)(\lambda^2-4\lambda+13)(\lambda-1)^4$
(c) $p_{\boldsymbol{A}}(\lambda) = (\lambda+5)(\lambda+2)^3(\lambda^2+6\lambda+25)^2$
(d) $p_{\boldsymbol{A}}(\lambda) = \lambda(\lambda^2+9)(\lambda^2+25)(\lambda-8)$

Exercise 1.13.4. Suppose that $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, and let $p_{\boldsymbol{A}}(\lambda) = \det(\boldsymbol{A}-\lambda\boldsymbol{I}_n)$ be the characteristic polynomial.

(a) If $n = 2$, show that $p_{\boldsymbol{A}}(\lambda)$ is a polynomial of degree two.
(b) If $n = 3$, show that $p_{\boldsymbol{A}}(\lambda)$ is a polynomial of degree three.
(c) Show that $p_{\boldsymbol{A}}(\lambda)$ is a polynomial of degree n.

Exercise 1.13.5. For each of the following matrices, write the vector $\boldsymbol{x} = (4\ -3)^{\mathrm{T}}$ as a linear combination of the eigenvectors. Explicitly give the weights (Fourier coefficients).

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}$
(b) $\boldsymbol{A} = \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}$
(c) $\boldsymbol{A} = \begin{pmatrix} 2 & -6 \\ 3 & -7 \end{pmatrix}$
(d) $\boldsymbol{A} = \begin{pmatrix} 1 & 4 \\ -2 & -3 \end{pmatrix}$

Exercise 1.13.6. Let $\boldsymbol{x} = (-3\ 5)^{\mathrm{T}}$. For each of the following matrices, write the vector $\boldsymbol{A}^{13}\boldsymbol{x}$ as a linear combination of the eigenvectors. Explicitly give the weights (Fourier coefficients). (*Hint*: Use Exercise 1.13.1)

(a) $\boldsymbol{A} = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}$
(b) $\boldsymbol{A} = \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}$
(c) $\boldsymbol{A} = \begin{pmatrix} 2 & -6 \\ 3 & -7 \end{pmatrix}$
(d) $\boldsymbol{A} = \begin{pmatrix} 1 & 4 \\ -2 & -3 \end{pmatrix}$

Exercise 1.13.7. Let $\boldsymbol{x} = (5\ 2\ -7)^{\mathrm{T}}$. For each of the following matrices, write the vector $\boldsymbol{A}^9\boldsymbol{x}$ as a linear combination of the eigenvectors. Explicitly give the weights (Fourier coefficients). (*Hint*: Use Exercise 1.13.1)

(a) $\boldsymbol{A} = \begin{pmatrix} 2 & 5 & -2 \\ 5 & 2 & 1 \\ 0 & 0 & -3 \end{pmatrix}$
(b) $\boldsymbol{A} = \begin{pmatrix} 3 & 0 & 0 \\ 1 & -1 & 2 \\ -2 & -2 & -1 \end{pmatrix}$

Exercise 1.13.8. Determine if each of the following statements is true or false. Provide an explanation for your answer.

(a) It is possible for $\boldsymbol{A} \in \mathbb{R}^{4\times 4}$ to have five eigenvalues.
(b) Every $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$ has two real eigenvalues.
(c) If $\boldsymbol{A} \in \mathbb{R}^{6\times 6}$, then $\boldsymbol{A}$ has at most six linearly independent eigenvectors.
(d) If $\boldsymbol{Ax} = \boldsymbol{0}$ has an infinite number of solutions, then all of the eigenvalues for $\boldsymbol{A}$ are nonzero.
(e) If $\boldsymbol{A} \in \mathbb{R}^{5\times 5}$, then it is possible for the characteristic polynomial to be of degree four.

1.14 Case studies

We now consider three problems in which it soon becomes clear that knowing the eigenvalues and associated eigenvectors for a given matrix clearly helps in understanding the solution.

1.14.1 Voter registration

Consider the following table:

	D	R	I
D	0.90	0.03	0.10
R	0.02	0.85	0.20
I	0.08	0.12	0.70

Here R represents Republicans, D Democrats, and I Independents. Let D_j, R_j, I_j represent the number of voters in each group in year j. The table provides information regarding the manner in which voters change their political affiliation from one year to the next. For example, reading down the first column we see that from one year to the next 90% of the Democrats remain Democrats, 2% become Republicans, and 8% become Independents. On the other hand, reading across the first row we see that the number of Democrats in a following year is the sum of 90% of the Democrats, 3% of the Republicans, and 10% of the Independents in the preceding year.

We wish to know what is the distribution of voters amongst the three groups after many years. Using the table we see that the number of voters in each group in year $n+1$ given the number of voters in each group in year n follows the rule,

$$\begin{aligned} D_{n+1} &= 0.90D_n + 0.03R_n + 0.10I_n \\ R_{n+1} &= 0.02D_n + 0.85R_n + 0.20I_n \\ I_{n+1} &= 0.08D_n + 0.12R_n + 0.70I_n. \end{aligned}$$

We implicitly assume here that the total number of voters is constant from one year to the next, so $D_n + R_n + I_n = N$ for any n, where N is the total number of voters. Upon setting

$$\boldsymbol{x}_n = \begin{pmatrix} D_n \\ R_n \\ I_n \end{pmatrix}, \quad \boldsymbol{M} = \begin{pmatrix} 0.90 & 0.03 & 0.10 \\ 0.02 & 0.85 & 0.20 \\ 0.08 & 0.12 & 0.70 \end{pmatrix},$$

we can rewrite this as the discrete *dynamical system*

$$\boldsymbol{x}_{n+1} = \boldsymbol{M}\boldsymbol{x}_n, \quad \boldsymbol{x}_0 \text{ given.} \tag{1.14.1}$$

The dynamical system (1.14.1) is known as a *Markov process*, and it is distinguished by the fact that the sum of each column of the transition (stochastic, Markov) matrix $\boldsymbol{M}$ is 1.

For a given initial distribution of voters $\boldsymbol{x}_0$, we wish to determine the distribution of voters after many years, i.e., we wish to compute $\lim_{n\to+\infty} \boldsymbol{x}_n$. First, we need to solve for $\boldsymbol{x}_n$. Since

$$\boldsymbol{x}_1 = \boldsymbol{M}\boldsymbol{x}_0, \quad \boldsymbol{x}_2 = \boldsymbol{M}\boldsymbol{x}_1 = \boldsymbol{M}(\boldsymbol{M}\boldsymbol{x}_0),$$

by defining $\boldsymbol{M}^k := \boldsymbol{M}\boldsymbol{M}\cdots\boldsymbol{M}$, i.e., $\boldsymbol{M}^k$ is the matrix $\boldsymbol{M}$ multiplied by itself k times, we have

$$\boldsymbol{x}_2 = \boldsymbol{M}^2\boldsymbol{x}_0.$$

Continuing in this fashion gives

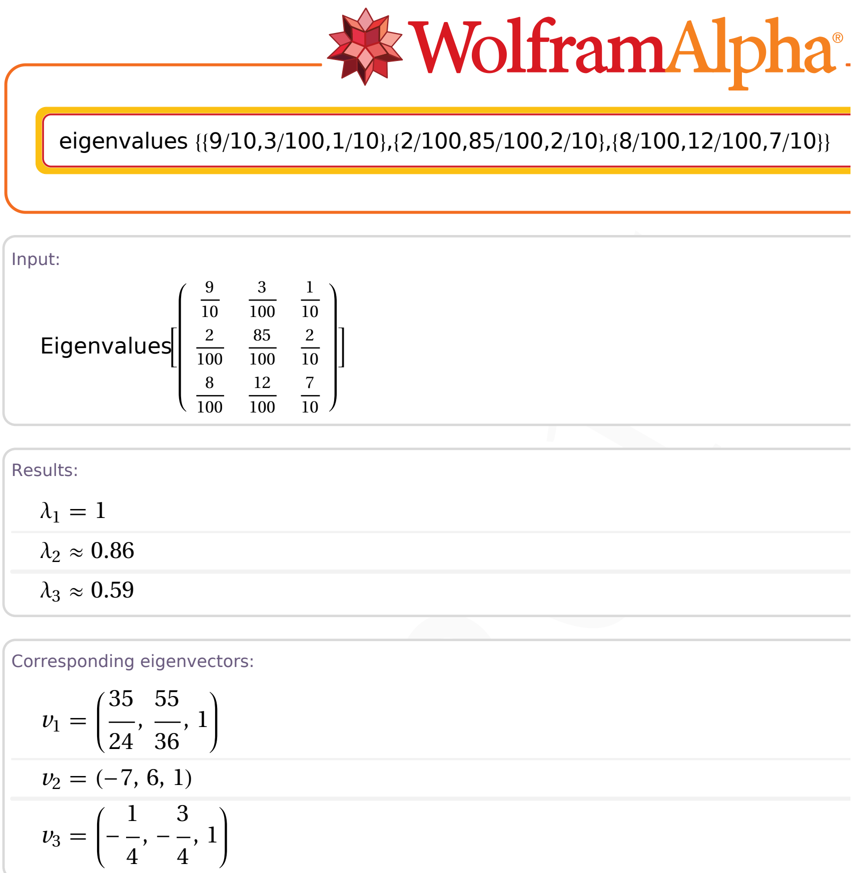
$$\boldsymbol{x}_3 = \boldsymbol{M}\boldsymbol{x}_2 = \boldsymbol{M}(\boldsymbol{M}^2\boldsymbol{x}_0) = \boldsymbol{M}^3\boldsymbol{x}_0, \quad \boldsymbol{x}_4 = \boldsymbol{M}\boldsymbol{x}_3 = \boldsymbol{M}(\boldsymbol{M}^3\boldsymbol{x}_0) = \boldsymbol{M}^4\boldsymbol{x}_0,$$

so by an induction argument that the solution to the dynamical system is

$$\boldsymbol{x}_n = \boldsymbol{M}^n\boldsymbol{x}_0. \tag{1.14.2}$$

Thus, our question is answered by determining $\lim_{n\to+\infty} \boldsymbol{M}^n$.

We now use the eigenvalues and eigenvectors of $\boldsymbol{M}$, and the Fourier expansion result of Lemma 1.13.9, in order to simplify the expression (1.14.2). These quantities are computed using WolframAlpha using exact arithmetic



Since the eigenvalues,

$$\lambda_1 = 1, \quad \lambda_2 \sim 0.86, \quad \lambda_3 \sim 0.59 \tag{1.14.3}$$

are distinct, by Theorem 1.13.8(d) the associated eigenvectors are linearly independent. Letting $\boldsymbol{P} = (\boldsymbol{v}_1\, \boldsymbol{v}_2\, \boldsymbol{v}_3)$, we know by Lemma 1.13.9 that the initial condition has a Fourier expansion in terms of the eigenvectors,

$$\boldsymbol{x}_0 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3 = \boldsymbol{P}\boldsymbol{c} \quad \Rightarrow \quad \boldsymbol{c} = \boldsymbol{P}^{-1}\boldsymbol{x}_0. \tag{1.14.4}$$

Now that the initial condition has been written in terms of the eigenvectors we can rewrite the solution in terms of the eigenvalues and eigenvectors. Via the linearity of matrix/vector multiplication, and using the expansion (1.14.4), we have (1.14.2) can be rewritten as

$$\boldsymbol{x}_n = c_1\boldsymbol{M}^n\boldsymbol{v}_1 + c_2\boldsymbol{M}^n\boldsymbol{v}_2 + c_3\boldsymbol{M}^n\boldsymbol{v}_3. \tag{1.14.5}$$

Regarding the term $\boldsymbol{M}^n\boldsymbol{v}_\ell$, for each $\ell = 1, 2, 3$

$$\boldsymbol{M}\boldsymbol{v}_\ell = \lambda_\ell \boldsymbol{v}_\ell \quad \Rightarrow \quad \boldsymbol{M}^2\boldsymbol{v}_\ell = \boldsymbol{M}(\boldsymbol{M}\boldsymbol{v}_\ell) = \lambda_\ell \boldsymbol{M}\boldsymbol{v}_\ell = \lambda_\ell^2 \boldsymbol{v}_\ell,$$

which by an induction argument leads to

$$\boldsymbol{M}^n\boldsymbol{v}_\ell = \lambda_\ell^n \boldsymbol{v}_\ell$$

(see Exercise 1.13.1). Substitution of the above into (1.14.5) then gives the solution in the form

$$\boldsymbol{x}_n = c_1\lambda_1^n\boldsymbol{v}_1 + c_2\lambda_2^n\boldsymbol{v}_2 + c_3\lambda_3^n\boldsymbol{v}_3, \quad \boldsymbol{c} = \boldsymbol{P}^{-1}\boldsymbol{x}_0. \tag{1.14.6}$$

We are now ready to determine the asymptotic limit of the solution. Using the eigenvalues as described in (1.14.3) we have

$$\lim_{n\to+\infty}\lambda_1^n = 1, \quad \lim_{n\to+\infty}\lambda_2^n = \lim_{n\to+\infty}\lambda_3^n = 0.$$

Consequently, for the solution formula (1.14.6) we have the asymptotic limit,

$$\lim_{n\to+\infty}\boldsymbol{x}_n = c_1\boldsymbol{v}_1, \quad \boldsymbol{v}_1 = \begin{pmatrix} 35/24 \\ 55/36 \\ 1 \end{pmatrix}.$$

From this formula we see that it is important only to determine c_1. Since the total number of people must be constant must be the same for each n, in the limit the total number of people is the same as the beginning number of people, i.e.,

$$c_1\left(\frac{35}{24} + \frac{55}{36} + 1\right) = N \quad \Rightarrow \quad c_1 = \frac{72}{287}N.$$

This observation allows us to write

$$c_1\boldsymbol{v}_1 = \frac{N}{287}\begin{pmatrix} 105 \\ 110 \\ 72 \end{pmatrix} \sim N\begin{pmatrix} 0.37 \\ 0.38 \\ 0.25 \end{pmatrix}.$$

In conclusion,

$$\lim_{n\to+\infty}\boldsymbol{x}_n \sim N\begin{pmatrix} 0.37 \\ 0.38 \\ 0.25 \end{pmatrix},$$

so in the long run 37% of the voters are Democrats, 38% are Republicans, and 25% are Independents. Note that this final distribution of voters is independent of the initial distribution of voters.

What is “long run” in this case? Since

$$\lambda_2^n, \lambda_3^n < 10^{-4} \quad \Rightarrow \quad n \geq 62,$$

the terms $c_2\lambda_2^n\boldsymbol{v}_2$ and $c_3\lambda_3^n\boldsymbol{v}_3$ in the solution Fourier expansion (1.14.6) will be negligible for $n \geq 62$. Thus, for $n \geq 62$ the solution will essentially be the asymptotic limit, which means that after 62 years the distribution of voters will be for all intents and purposes that given above.

1.14.2 Discrete SIR model

The dynamics of epidemics are often based upon SIR models. In a given population there are three subgroups:

(a) susceptible (S): those who are able to get a disease, but have not yet been infected

(b) infected (I): those who are currently fighting the disease
(c) recovered (R): those who have had the disease, or are immune.

Although it is not necessary, it is often assumed that the entire population,

$$N = S + I + R,$$

is a constant. Moreover, it is assumed that the number of people in each group does not depend upon location. Consequently, the model to be given is reasonable when looking at epidemics in a school environment, but it not very good when trying to understand nationwide outbreaks of disease (which are generally more regional).

Paladini et al. [36] provide a descriptive discrete-time dynamical system of the form

$$\begin{aligned} S_{n+1} &= qS_n + cR_n \\ I_{n+1} &= (1-q)S_n + bI_n \\ R_{n+1} &= (1-b)I_n + (1-c)R_n. \end{aligned} \tag{1.14.7}$$

Here S_j is the number of susceptible people in the sampling interval j, I_j is the number of infected people in the sampling interval j, and R_j is the number of recovered people in the sampling interval j. Depending upon the disease being studied, the sampling interval may be monthly, yearly, or even larger. The model assumes that:

(a) susceptible people must become infected before recovering
(b) infected people must recover before again becoming susceptible
(c) recovered people cannot become infected without first becoming susceptible.

As for the parameters, we have:

(a) $0 \le q \le 1$ is the probability that a susceptible avoids infection
(b) $0 \le b \le 1$ is the proportion of individuals which remain infected
(c) $0 \le c \le 1$ is the fraction of recovered individuals which lose immunity.

The probability parameter q is generally a function of both S and I, e.g.,

$$q = 1 - p\frac{I}{N},$$

where p is the probability of the infection being transmitted through a time of contact. We will assume that q is fixed; in particular, we will assume that it does not depend upon the proportion of infected people. It is not difficult to check that

$$S_{n+1} + I_{n+1} + R_{n+1} = S_n + I_n + R_n,$$

so the total population remains constant for all n (see Exercise 1.14.3). We could use this fact to reduce the number of variables in (1.14.7), but we will not do so in our analysis.

We now proceed to solve (1.14.7). Upon setting

$$\boldsymbol{x}_n = \begin{pmatrix} S_n \\ I_n \\ R_n \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} q & 0 & c \\ 1-q & b & 0 \\ 0 & 1-b & 1-c \end{pmatrix},$$

we can rewrite the dynamical system in the form

$$\boldsymbol{x}_{n+1} = \boldsymbol{A}\boldsymbol{x}_n. \tag{1.14.8}$$

Note that this dynamical system shares (at least) one feature with the Markov process associated with the voter registration problem: the sum of each column of the matrix $\boldsymbol{A}$ is 1. By following the argument leading to (1.14.2) we know the solution is

$$\boldsymbol{x}_n = \boldsymbol{A}^n \boldsymbol{x}_0.$$

Moreover, we know that the eigenvalues and associated eigenvectors of $\boldsymbol{A}$ can be used to simplify the form of the solution. Writing $\boldsymbol{A}\boldsymbol{v}_j = \lambda_j \boldsymbol{v}_j$ for $j = 1,2,3$, we know that the solution can be written

$$\boldsymbol{x}_n = c_1 \lambda_1^n \boldsymbol{v}_1 + c_2 \lambda_2^n \boldsymbol{v}_2 + c_3 \lambda_3^n \boldsymbol{v}_3. \tag{1.14.9}$$

The underlying assumption leading to the solution formula in (1.14.9), which will be verified for specific values of b, c, q, is that the eigenvectors are linearly independent. We then know by Lemma 1.13.9 that the initial condition $\boldsymbol{x}_0$ has the Fourier expansion

$$\boldsymbol{x}_0 = c_1 \boldsymbol{v}_1 + c_2 \boldsymbol{v}_2 + c_3 \boldsymbol{v}_3,$$

and the Fourier coefficients are found by solving the linear system

$$\boldsymbol{P}\boldsymbol{c} = \boldsymbol{x}_0 \quad \Rightarrow \quad \boldsymbol{c} = \boldsymbol{P}^{-1}\boldsymbol{x}_0, \quad \boldsymbol{P} = (\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3).$$

We are now ready to determine the asymptotic limit of the solution. Following [36] we will assume

$$b = 0.5, \quad c = 0.01.$$

If we further assume $q = 0.2$, then the matrix $\boldsymbol{A}$ becomes

$$\boldsymbol{A} = \begin{pmatrix} 0.2 & 0 & 0.01 \\ 0.8 & 0.5 & 0 \\ 0 & 0.5 & 0.99 \end{pmatrix}.$$

Using WolframAlpha and exact arithmetic we find the eigenvalues and associated eigenvectors to be:

WolframAlpha

eigenvalues {{2/10,0,1/100},{8/10,5/10,0},{0,5/10,99/100}}

Input:

$$\text{Eigenvalues}\left[\begin{pmatrix} \frac{2}{10} & 0 & \frac{1}{100} \\ \frac{8}{10} & \frac{5}{10} & 0 \\ 0 & \frac{5}{10} & \frac{99}{100} \end{pmatrix}\right]$$

Results:

$\lambda_1 = 1$

$\lambda_2 \approx 0.47159$

$\lambda_3 \approx 0.21841$

Corresponding eigenvectors:

$$v_1 = \left(\frac{1}{80}, \frac{1}{50}, 1\right)$$

$$v_2 = \left(\frac{1}{100}\left(29 - \sqrt{641}\right), \frac{1}{100}\left(-129 + \sqrt{641}\right), 1\right)$$

$$v_3 = \left(\frac{1}{100}\left(29 + \sqrt{641}\right), \frac{1}{100}\left(-129 - \sqrt{641}\right), 1\right)$$

Because the eigenvalues are distinct,

$$\lambda_1 = 1, \quad \lambda_2 \sim 0.47, \quad \lambda_3 \sim 0.22,$$

we know that the associated eigenvectors are linearly independent.

Since

$$\lim_{n\to+\infty} \lambda_1^n = 1, \quad \lim_{n\to+\infty} \lambda_2^n = \lim_{n\to+\infty} \lambda_3^n = 0,$$

we have the asymptotic limit

$$\lim_{n\to+\infty} \boldsymbol{x}_n = c_1 \boldsymbol{v}_1, \quad \boldsymbol{v}_1 = \begin{pmatrix} 1/80 \\ 1/50 \\ 1 \end{pmatrix}.$$

We see that we now must determine c_1. Since the total number of people is constant for all n, in the limit the total number of people is the same as the beginning number of people, which leads to

$$c_1\left(\frac{1}{80} + \frac{1}{50} + 1\right) = N \quad \Rightarrow \quad c_1 = \frac{400}{413} N.$$

This observation allows us to write

$$c_1\boldsymbol{v}_1 = \frac{N}{413}\begin{pmatrix} 5 \\ 8 \\ 400 \end{pmatrix} \sim N\begin{pmatrix} 0.012 \\ 0.019 \\ 0.969 \end{pmatrix}.$$

In conclusion,

$$\lim_{n\to+\infty} \boldsymbol{x}_n = N\begin{pmatrix} 0.012 \\ 0.019 \\ 0.969 \end{pmatrix},$$

so in the long run 1.2% of the people are susceptible, 1.9% of the people are infected, and 96.9% of the people are recovered. Note that this final distribution of the population is independent of the number of people who were originally infected.

What is "long run" in this case? Since

$$\lambda_2^n,\, \lambda_3^n < 10^{-4} \quad\Rightarrow\quad n \ge 13,$$

the terms $c_2\lambda_2^n\boldsymbol{v}_2$ and $c_3\lambda_3^n\boldsymbol{v}_3$ in the solution Fourier expansion (1.14.6) will be negligible for $n \ge 13$. Thus, for $n \ge 13$ the solution will essentially be the asymptotic limit, which means that after 13 years the distribution of people will be for all intents and purposes that given above.

1.14.3 Northern spotted owl

The size of the Northern spotted owl population is closely associated with the health of the mature and old-growth coniferous forests in the Pacific Northwest. Over the last few decades there has been loss and fragmentation of these forests, which may potentially effect the long-term survival of this species of owl. For spotted owls there are three distinct groupings:

(a) juveniles (j) under one year old
(b) subadults (s) between one and two years old
(c) adults (a) two years old and older.

The owls mate during the latter two life stages, and begin breeding as adults.

In year n let j_n be the number of juveniles, s_n be the number of subadults, and a_n be the number of adults. Mathematical ecologists have modeled a particular spotted owl population via the discrete dynamical system

$$\begin{aligned} j_{n+1} &= 0.33a_n \\ s_{n+1} &= 0.18j_n \\ a_{n+1} &= 0.71s_n + 0.94a_n. \end{aligned}$$

The juvenile population in the next year is 33% of the adult population, 18% of the juveniles in one year become subadults in the next year, 71% of the subadults in one year become adults the next year, and 94% of the adults survive from one year to the next (see Lamberson et al. [25], Lay [26, Chapter 5], and the references therein). Upon setting

$$\boldsymbol{x}_n = \begin{pmatrix} j_n \\ s_n \\ a_n \end{pmatrix},$$

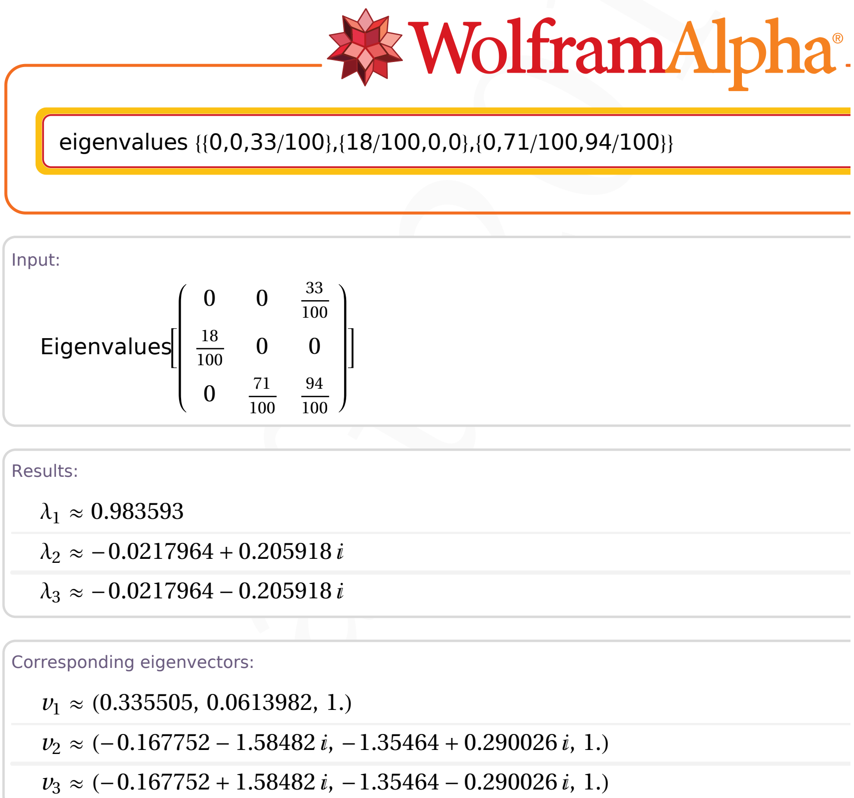
we can rewrite this dynamical system in the form

$$\boldsymbol{x}_{n+1} = \boldsymbol{A}\boldsymbol{x}_n, \quad \boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0.33 \\ 0.18 & 0 & 0 \\ 0 & 0.71 & 0.94 \end{pmatrix}. \tag{1.14.10}$$

For a given initial distribution of owls, we wish to see what is the distribution of the owls after many years. We first solve for $\boldsymbol{x}_n$ in terms of $\boldsymbol{x}_0$. Following the argument leading to (1.14.2) we know the solution is

$$\boldsymbol{x}_n = \boldsymbol{A}^n \boldsymbol{x}_0. \tag{1.14.11}$$

Following our discussion in the previous case study, we know that we next wish to use the eigenvalues and associated eigenvectors for the matrix $\boldsymbol{A}$. These quantities are computed using WolframAlpha using exact arithmetic (the exact expressions are not shown)



In particular, the eigenvalues are

$$\lambda_1 \sim 0.98, \quad \lambda_2 \sim -0.02 + \mathrm{i}0.21, \quad \lambda_3 \sim -0.02 - \mathrm{i}0.21.$$

By following the logic leading to (1.14.6) we know the solution is

$$\boldsymbol{x}_n = c_1\lambda_1^n \boldsymbol{v}_1 + c_2\lambda_2^n \boldsymbol{v}_2 + c_3\lambda_3^n \boldsymbol{v}_3, \quad \boldsymbol{x}_0 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3. \tag{1.14.12}$$

Because the eigenvectors are complex-valued, the Fourier coefficients may also be.

The asymptotic behavior of the solution depends on the size of the eigenvalues. Looking back to the solution formula (1.14.12), we need to understand what happens when we take successive powers of the eigenvalues. While we understand what λ_1^n means, we do not have an intuitive understanding as to what it means when we write λ_2^n and λ_3^n. Recall that we showed in (1.12.1) complex numbers $z = a + \mathrm{i}b$ can be written in the polar form

$$z = |z|\mathrm{e}^{\mathrm{i}\theta}; \quad \mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta,\ \tan\theta = \frac{b}{a}.$$

The polar representation of complex numbers allows us to write

$$z^n = |z|^n \left(\mathrm{e}^{\mathrm{i}\theta}\right)^n = |z|^n \mathrm{e}^{\mathrm{i}n\theta}.$$

In particular, we have

$$\begin{aligned} \lambda_2 = 0.21\mathrm{e}^{\mathrm{i}0.53\pi} \quad &\Rightarrow \quad \lambda_2^n = 0.21^n \mathrm{e}^{\mathrm{i}0.53n\pi} \\ \lambda_3 = 0.21\mathrm{e}^{-\mathrm{i}0.53\pi} \quad &\Rightarrow \quad \lambda_3^n = 0.21^n \mathrm{e}^{-\mathrm{i}0.53n\pi}. \end{aligned}$$

Since

$$|\mathrm{e}^{\mathrm{i}\theta}| = 1 \quad \Rightarrow \quad |\mathrm{e}^{\mathrm{i}n\theta}| = |\mathrm{e}^{\mathrm{i}\theta}|^n = 1,$$

the magnitude of z^n is controlled solely by the magnitude of z,

$$|z^n| = |z|^n \left|\mathrm{e}^{\mathrm{i}n\theta}\right| = |z|^n.$$

Thus, in our example it will be the case that $|\lambda_2^n| = |\lambda_3^n| < 10^{-4}$ for $n \ge 6$. Going back to the solution formula (1.14.12), we see that for $n \ge 6$ we can write it as

$$\boldsymbol{x}_n \sim 0.98^n c_1 \boldsymbol{v}_1.$$

In order to properly interpret this solution formula, we first want to write the eigenvector $\boldsymbol{v}_1$ so that each entry corresponds to the percentage of the total owl population in each subgroup. This requires that the entries of the eigenvector sum to one. Unfortunately, WolframAlpha does not present an eigenvector with that property; instead, we get

$$\boldsymbol{v}_1 \sim \begin{pmatrix} 0.3355 \\ 0.0613 \\ 1.0000 \end{pmatrix}.$$

We know that eigenvectors are not unique, and can be scaled in any desired fashion. We find the rescaled version of the eigenvector using WolframAlpha

WolframAlpha

{0.3355,0.0613,1.0}/(0.3355+0.0613+1.0)

Input:

$$\frac{\{0.3355,\ 0.0613,\ 1.\}}{0.3355+0.0613+1.}$$

Result:

{0.240192, 0.043886, 0.715922}

The desired eigenvector is approximately

$$\boldsymbol{v}_1 \sim \begin{pmatrix} 0.24 \\ 0.04 \\ 0.72 \end{pmatrix}.$$

As for the constant c_1, we first rewrite the system (1.14.12) in matrix/vector form,

$$\boldsymbol{x}_0 = \boldsymbol{P}\boldsymbol{c}, \quad \boldsymbol{P} = (\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3).$$

We numerically compute $\boldsymbol{P}^{-1}$ using WolframAlpha

WolframAlpha

inverse {{0.2402,–0.1678–1.5848i,–0.1678+1.5848i},{0.0439,–1.3546+0.2900i,–1.

Input interpretation:

$$\begin{pmatrix} 0.2402 & -0.1678-1.5848\,i & -0.1678+1.5848\,i \\ 0.0439 & -1.3546+0.29\,i & -1.3546-0.29\,i \\ 0.7159 & 1. & 1. \end{pmatrix}^{-1} \quad \text{(matrix inverse)}$$

i is

Result:

$$\begin{pmatrix} 0.169497+0.\,i & 0.926274+5.55112\times10^{-17}\,i & 1.28317-3.40402\times10^{-18}\,i \\ -0.0606716+0.296228\,i & -0.33156-0.105301\,i & 0.0406884-0.0929338\,i \\ -0.0606716-0.296228\,i & -0.33156+0.105301\,i & 0.0406884+0.0929338\,i \end{pmatrix}$$

This is a numerical calculation, so the first row of the inverse is actually composed of purely real numbers. Upon writing $\boldsymbol{x}_0 = (j_0\ s_0\ a_0)^{\mathrm{T}}$, we have

$$\boldsymbol{c} = \boldsymbol{P}^{-1}\boldsymbol{x}_0 \quad \Rightarrow \quad c_1 \sim 0.17 j_0 + 0.93 s_0 + 1.28 a_0.$$

In conclusion, we have

$$c_1\boldsymbol{v}_1 \sim [0.17j_0 + 0.93s_0 + 1.28a_0]\begin{pmatrix}0.24\\0.04\\0.72\end{pmatrix}.$$

Thus, for $n \geq 6$ we can say

$$\boldsymbol{x}_n \sim 0.98^n[0.17j_0 + 0.93s_0 + 1.28a_0]\begin{pmatrix}0.24\\0.04\\0.72\end{pmatrix}.$$

Roughly 24% of the owls will be juveniles, 4% of the owls will be subadults, and 72% of the owls will be adults. The total number of owls in each group will depend on the initial distribution. The overall population will slowly decrease, and assuming no changes in the conditions leading to the original model (1.14.10) the owls will *eventually* become extinct (e.g., $0.98^n \leq 0.1$ for $n \geq 114$).

Exercises

Exercise 1.14.1. Consider the below table, which represents the fraction of the population in each group - City (C), and Suburban (S) - which migrates to a different group in a given year. Assume that the total population is constant. Further assume that there are initially 1500 city dwellers, and 1000 suburbanites.

(a) How many people will there be in each group after many years? Assume that the total number of people is constant.
(b) Does your answer in (a) depend on the initial number of people in each group?

	C	S
C	0.94	0.15
S	0.06	0.85

Exercise 1.14.2. Consider the below table, which represents the fraction of the population in each group - City (C), Suburban (S), and Rural (R) - which migrates to a different group in a given year. Assume that the total population is constant. Further assume that there are initially 1000 city dwellers, 750 suburbanites, and 250 rural dwellers.

(a) How many people will there be in each group after many years? Assume that the total number of people is constant.
(b) Does your answer in (a) depend on the initial number of people in each group?

	C	S	R
C	0.91	0.09	0.02
S	0.05	0.87	0.08
R	0.04	0.04	0.90

Exercise 1.14.3. Consider the SIR model given in (1.14.7).

(a) Show that the model supports that the total population is fixed, i.e., show that $S_n + I_n + R_n = S_0 + I_0 + R_0$ for all $n \geq 1$.

(b) Writing $N = S_n + I_n + R_n$, show that the system is equivalent to

$$\begin{aligned} S_{n+1} &= (q-1)S_n - cI_n + cN \\ I_{n+1} &= (1-q)S_n + bI_n. \end{aligned}$$

(c) If one solves for (S_n, I_n) in part (b), how is R_n found?

Exercise 1.14.4. Consider the SIR case study of Chapter 1.14.2. Suppose that $b = 0.8$ and $c = 0.01$. Further suppose that n is large.

(a) If $q = 0.2$, what percentage of the total population will be comprised of infected people?
(b) If $q = 0.7$, what percentage of the total population will be comprised of infected people?

Exercise 1.14.5. Consider the SIR case study of Chapter 1.14.2. Suppose that $c = 0.1$ and $q = 0.4$. Further suppose that n is large.

(a) If $b = 0.05$, what percentage of the total population will be comprised of recovered people?
(b) If $b = 0.35$, what percentage of the total population will be comprised of recovered people?

Exercise 1.14.6. Consider the SIR case study of Chapter 1.14.2. Suppose that $b = 0.4$ and $q = 0.3$. Further suppose that n is large.

(a) If $c = 0.02$, what percentage of the total population will be comprised of susceptible people?
(b) If $c = 0.25$, what percentage of the total population will be comprised of susceptible people?

Exercise 1.14.7. Consider the case study of the Northern spotted owl in Chapter 1.14.3. Let the fraction of subadults who become adults be represented by the parameter r (replace 0.71 with r in the matrix of (1.14.10)). Suppose that n is large.

(a) If $r = 0.1$, what percentage of the total population will be comprised of subadults?
(b) If $r = 0.45$, what percentage of the total population will be comprised of subadults?

Exercise 1.14.8. Consider the case study of the Northern spotted owl in Chapter 1.14.3. Let the fraction of adults who survive from one year to the next be represented by the parameter r (replace 0.94 with r in the matrix of (1.14.10)). Suppose that n is large.

(a) If $r = 0.65$, what percentage of the total population will be comprised of adults?
(b) If $r = 0.85$, what percentage of the total population will be comprised of adults?

Group projects

1.1. A second-order discrete dynamical system which arises, e.g., when numerically solving a *Sturm-Liouville problem*, is

$$y_{j+2} - 2y_{j+1} + y_j = \lambda y_j.$$

Here λ is a real-valued parameter. Sturm-Liouville problems arise in the study of vibrating strings, heat flow in an insulated wire, and quantum mechanics. The boundary conditions to be considered here are

$$y_0 = y_{n+1} = 0.$$

One solution to this problem is the trivial solution, $y_j = 0$ for $j = 1, \dots, n$, and it is valid for any value of λ. The goal is find those values of λ for which there is a nontrivial solution.

(a) Setting

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n,$$

find the symmetric matrix $\boldsymbol{A}$ so that the problem can be rewritten

$$\boldsymbol{A}\boldsymbol{y} = \lambda\boldsymbol{y}.$$

(b) Suppose that $n = 2$. Find the eigenvalues and associated eigenvectors.

The remaining questions are devoted to finding an explicit expression for the eigenvalues if $n \geq 3$.

(c) Set $h = 1/(n+1)$, and for an as yet unknown α write

$$y_j = \mathrm{e}^{\mathrm{i}\alpha jh} = \cos(\alpha jh) + \mathrm{i}\sin(\alpha jh).$$

By using the original system show that

$$\mathrm{e}^{\mathrm{i}\alpha(j+1)h} + \mathrm{e}^{\mathrm{i}\alpha(j-1)h} = (2+\lambda)\mathrm{e}^{\mathrm{i}\alpha jh}.$$

(d) Solve the equation in part (c) and show that

$$\lambda = 2(1 - \cos(\alpha h)).$$

(e) The solution as written is complex-valued. Use linearity to show that two possible real-valued solutions are

$$y_j = \cos(\alpha jh) \quad \text{or} \quad y_j = \sin(\alpha jh).$$

(f) Use the boundary condition at $j = 0$, $y_0 = 0$, to reject the first solution and conclude,

$$y_j = \sin(\alpha jh), \quad j = 0, \dots, n+1.$$

(g) Use the boundary condition at $j = n+1$, $y_{n+1} = 0$, to determine the values of α which lead to a nontrivial solution.

(h) What are the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ for the matrix $\boldsymbol{A}$?

1.2. Let $\boldsymbol{S} \in \mathbb{R}^{n\times n}$ be a symmetric matrix, $\boldsymbol{S}^{\mathrm{T}} = \boldsymbol{S}$. We will suppose that all of the eigenvalues are distinct. While we will not prove it, the eigenvalues must be real; hence, we can order them,

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n.$$

The associated eigenvectors will be denoted $\boldsymbol{v}_j$ for $j = 1,\ldots,n$. Recall the definition of an inner-product given in Definition 1.8.1.

(a) Show that $\langle \boldsymbol{Sx}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{Sy}\rangle$.

(b) Show that the associated eigenvectors are orthogonal, $\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = 0$ for $i \neq j$. It may be useful to:

- consider the quantity $\langle \boldsymbol{Sv}_i, \boldsymbol{v}_j\rangle$
- use part (a).

(c) From part (b) we can normalize the eigenvectors to be orthonormal,

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

If for a given vector $\boldsymbol{x}$ we write the Fourier expansion,

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n,$$

what are the coefficients c_j for $j = 1,\ldots,n$?

(d) What is the Fourier expansion for $\boldsymbol{Sx}$? How do the coefficients for $\boldsymbol{Sx}$ relate to those of $\boldsymbol{x}$?

(e) If $\lambda_1 > 0$, i.e., if all of the eigenvalues are positive, show that $\langle \boldsymbol{Sx}, \boldsymbol{x}\rangle > 0$ for any vector $\boldsymbol{x}$.

(f) If $\lambda_n < 0$, i.e., if all of the eigenvalues are negative, show that $\langle \boldsymbol{Sx}, \boldsymbol{x}\rangle < 0$ for any vector $\boldsymbol{x}$.

1.3. A *state-space model* of a control system includes a difference equation of the form

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{u}_k, \quad k = 0, 1, \ldots.$$

The vector $\boldsymbol{x}_k$ is the state vector at "time" k, and the vector $\boldsymbol{u}_k$ is a control or input at time k. The size of the matrices is $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n\times m}$, so $\boldsymbol{x}_k \in \mathbb{R}^n$ and $\boldsymbol{u}_k \in \mathbb{R}^m$. The pair $(\boldsymbol{A}, \boldsymbol{B})$ is said to be *controllable* if for a given $\boldsymbol{x}^*$ there is a sequence of control vectors $\{\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-1}\}$ such that $\boldsymbol{x}_n = \boldsymbol{x}^*$.

(a) For a given initial state $\boldsymbol{x}_0$, find the state:

- $\boldsymbol{x}_1$
- $\boldsymbol{x}_2$
- $\boldsymbol{x}_n$ for $n \geq 3$.

Write your solution only in terms of the initial state $\boldsymbol{x}_0$ and the control vectors.

(b) In block-matrix notation we can write

$$\boldsymbol{C}_0\boldsymbol{a}_0 + \boldsymbol{C}_1\boldsymbol{a}_1 = (\boldsymbol{C}_0\ \boldsymbol{C}_1)\begin{pmatrix} \boldsymbol{a}_0 \\ \boldsymbol{a}_1 \end{pmatrix}.$$

Use this notation to rewrite your solution formula for $\boldsymbol{x}_n$ in the form,

$$\boldsymbol{x}_n = \boldsymbol{C}_0\boldsymbol{x}_0 + \boldsymbol{C}_1\boldsymbol{y},$$

for some matrices $\boldsymbol{C}_0$ and $\boldsymbol{C}_1$, and some vector $\boldsymbol{y}$ which depends upon the input sequence.

(c) Show that if

$$\operatorname{rank}\left(\boldsymbol{A}^{n-1}\boldsymbol{B}\ \boldsymbol{A}_{n-2}\boldsymbol{B}\ \cdots\ \boldsymbol{A}\boldsymbol{B}\ \boldsymbol{B}\right) = n,$$

then the system is controllable.

(d) If

$$\boldsymbol{A} = \begin{pmatrix} 0.8 & -0.3 & 0.0 \\ 0.2 & 0.5 & 1.0 \\ 0.0 & 0.0 & -0.5 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

determine if the pair $(\boldsymbol{A}, \boldsymbol{B})$ is controllable.