

Notes for Math 152: Linear Systems

Spring, 2017

Richard Froese and Brian Wetton

© 2016 Richard Froese and Brian Wetton. Permission is granted to make and distribute copies of this document provided the copyright notice and this permission notice are preserved on all copies.

Contents

1	Introduction	5
1.1	Course Goals	5
1.2	About the Subject	6
1.2.1	Connection to Geometry	6
1.2.2	Linear Systems	6
1.2.3	Eigen-analysis	7
1.3	About The Computer Labs	8
1.4	About These Notes	9
2	Vectors and Geometry	11
2.1	Chapter Introduction	11
2.2	Vectors	11
2.2.1	Multiplication by a number and vector addition	12
2.2.2	Co-ordinates	13
2.2.3	Properties of vector addition and scalar multiplication	17
2.2.4	MATLAB: basic scalar and vector operations	18
2.2.5	Problems	19
2.3	Geometrical Aspects of Vectors	20
2.3.1	Length of a vector	20
2.3.2	The dot product	20
2.3.3	Projections and Unit Vectors	23
2.3.4	MATLAB: <code>norm</code> and <code>dot</code> commands	25
2.3.5	Problems	25
2.4	Determinants and the Cross Product	27
2.4.1	The determinant in two and three dimensions	27
2.4.2	The cross product	29
2.4.3	The triple product and the determinant in three dimensions	32
2.4.4	MATLAB: assigning matrices and <code>det</code> and <code>cross</code> commands	33
2.4.5	MATLAB: generating scripts with the MATLAB editor	34
2.4.6	MATLAB: floating point representation of real numbers	35
2.4.7	Problems	36
2.5	Lines and Planes	37
2.5.1	Describing linear sets	37
2.5.2	Lines in two dimensions: Parametric form	37

2.5.3	Lines in two dimensions: Equation form	38
2.5.4	Lines in three dimensions: Parametric form	39
2.5.5	Lines in three dimensions: Equation form	40
2.5.6	Planes in three dimensions: Parametric form	41
2.5.7	Planes in three dimensions: Equation form	42
2.5.8	Problems	43
2.6	Introduction to Linear Systems	45
2.6.1	Description of points and the geometry of solutions to systems of equations	45
2.6.2	Describing the whole plane in two dimensions and all of space in three dimensions	48
2.6.3	Linear dependence and independence	49
2.6.4	Problems	51
2.7	Additional Topics	52
2.7.1	Application: rotational motion	52
2.7.2	Application: 3-D graphics	53
2.8	Solutions to Chapter Problems	56
3	Solving Linear Systems	70
3.1	Linear Systems	70
3.1.1	General Form of Linear Systems	70
3.1.2	Solving Linear Systems by Substitution	70
3.1.3	Elementary row (equation) operations	71
3.1.4	Augmented Matrices	72
3.1.5	Problems	73
3.2	Gaussian Elimination	74
3.2.1	Using MATLAB for row reductions	80
3.2.2	Problems	81
3.3	Homogeneous Equations	83
3.3.1	Properties of solutions of homogeneous systems.	85
3.3.2	Connection of solutions to homogeneous and inhomogeneous systems.	85
3.3.3	Problems	87
3.4	Geometric Applications	87
3.4.1	Problems	89
3.5	Resistor Networks	90
3.5.1	Elements of Basic Circuits	90
3.5.2	Two Simple Examples Made Complicated	91
3.5.3	Loop Currents	93
3.5.4	Alternate Presentation of Resistor Networks	97
3.5.5	Problems	101
3.6	Additional Topics	104
3.6.1	The Check Column	104
3.6.2	Quadratic Functions	106
3.6.3	Least squares fit	108
3.6.4	Equilibrium configuration of hanging weights and springs	109

3.6.5	Problems	111
3.7	Solutions to Chapter Problems	111
4	Matrices and Determinants	126
4.1	Matrix operations	126
4.1.1	MATLAB	129
4.1.2	Problems	129
4.2	Linear Transformations and Matrices	130
4.2.1	Linear Transformations	130
4.2.2	Rotations in two dimensions	131
4.2.3	Projections in two dimensions	132
4.2.4	Reflections in two dimensions	133
4.2.5	Every linear transformation is multiplication by a matrix	134
4.2.6	Composition of linear transformations and matrix product	136
4.2.7	Problems	136
4.3	Application: random walks	139
4.3.1	Problems	144
4.4	The Transpose	146
4.4.1	MATLAB	147
4.4.2	Problems	147
4.5	Matrix Inverses	147
4.5.1	Computing the inverse	150
4.5.2	Inverses of Products	153
4.5.3	MATLAB	154
4.5.4	Problems	154
4.6	Determinants	155
4.6.1	Definition of Determinants	155
4.6.2	Determinants of Triangular matrices	156
4.6.3	Summary of determinant calculation rules	157
4.6.4	Calculation of determinant using row operations	158
4.6.5	More expansion formulae	158
4.6.6	MATLAB	160
4.6.7	Problems	160
4.7	Additional Topics	162
4.7.1	Return to Resistor Networks	162
4.7.2	Application: General Least Squares	166
4.7.3	Least squares solutions	166
4.7.4	Problems	167
4.7.5	Elementary matrices	168
4.7.6	Problems	171
4.7.7	Exchanging two rows changes the sign of the determinant	171
4.7.8	The determinant is linear in each row separately	173
4.7.9	Adding a multiple of one row to another doesn't change the determinant	174
4.7.10	The determinant of QA	175
4.7.11	The determinant of A is zero exactly when A is not invertible	176

4.7.12	The product formula: $\det(AB) = \det(A)\det(B)$	176
4.7.13	The determinant of the transpose	177
4.7.14	An impractical formula for the inverse	177
4.7.15	Cramer's rule, an impractical way to solve systems	178
4.7.16	Problems	179
4.8	Solutions to Chapter Problems	179
5	Complex numbers	202
5.1	Complex arithmetic	202
5.2	Complex matrices and linear systems	203
5.2.1	Determinants of complex matrices	204
5.2.2	Homogeneous complex linear systems	204
5.2.3	Non-homogeneous complex linear systems	206
5.2.4	Finding inverses of complex matrices	206
5.3	Complex exponential	206
5.4	Polar representation of a complex number	208
5.5	MATLAB	208
5.6	Problems	209
5.7	Solutions to Chapter Problems	209
6	Eigen-analysis	212
6.1	Eigenvalues and eigenvectors	212
6.1.1	Computing the eigenvalues and eigenvectors	213
6.1.2	Complex eigenvalues and eigenvectors	219
6.1.3	MATLAB	221
6.1.4	Problems	222
6.2	Eigenanalysis simplifies matrix powers	223
6.2.1	Problems	227
6.3	Systems of linear differential equations	227
6.3.1	Problems	233
6.4	LCR circuits	234
6.4.1	Capacitors and inductors	234
6.4.2	Differential equations for LCR circuits	235
6.4.3	Alternate description of LCR circuits	237
6.4.4	Problems	239
6.5	Additional Topics	239
6.5.1	Diagonalization	239
6.5.2	Computing high powers of a matrix	241
6.5.3	Another formula for the determinant	242
6.5.4	The matrix exponential and differential equations	242
6.5.5	Converting higher order equations into first order systems	243
6.5.6	Springs and weights	245
6.5.7	Problems	250
6.6	Solutions to Chapter Problems	251

Chapter 1

Introduction

Linear Algebra is a branch of Mathematics. Many applied problems from Science, Engineering and Finance can be written in terms of Linear Algebra questions. This is also true of Calculus, which is why these two fields are stressed in undergraduate Mathematics education at UBC and other universities. By putting a class of commonly occurring problems into a unified, abstract framework, the problems can be studied in detail and well understood. This understanding can then be taken back to problems in many different fields. It is especially important to realize that very large problems (with potentially millions of unknowns) can be understood in the same framework as the model problems we do by hand in these courses. This course has computer labs that involve the mathematical software, MATLAB, that will allow you to solve larger problems using numerical computations.

Unlike Calculus, Linear Algebra does not require a lot of background knowledge. Basic operations in Linear Algebra are just arithmetic. However, there is a powerful connection between these simple arithmetic operations and geometric quantities. Simple ideas in this course start to become abstract when combined together. Linear Algebra is a subject you can study with very limited mathematical background, but you are advised to keep up with course lectures, readings, assignments and computer labs so you won't be left behind at the transition from concrete ideas to abstract ones.

1.1 Course Goals

The goal of the course is to enable students to

1. recognize linear algebra questions (for which there are straight-forward analytic and numerical solution techniques) as parts of applied problems
2. make the connection between geometric properties and analytic quantities (determinants, dot and cross products, eigenvalues, etc.)

3. recognize that linear systems of equations can have unique, infinite or no solutions and know how to determine all solutions or that none exist
4. recognize matrix multiplication as a linear transformation and that such transformations (to the same dimensional space) can be simplified using eigen-analysis
5. use complex numbers, which arise naturally in the eigen-analysis of matrices

1.2 About the Subject

The subject of the course is Linear Algebra, focussing on four main topics:

1. vectors and matrices and connections to geometry
2. linear systems
3. complex numbers
4. eigen-analysis of matrices

Several applications are considered. The two applications that come up repeatedly in the notes are resistor networks and random walks.

1.2.1 Connection to Geometry

The first topic considered in the course is vectors, which are quantities with both magnitude and direction. A typical quantity represented as a vector is a force \mathbf{F} on an object as shown in Figure 1.1. The force \mathbf{F} in that figure acts in the $x - y$ plane as shown. The vector force \mathbf{F} can be represented by its components (F_x, F_y) . Some interesting questions you will be able to answer after completing this course are: what directions are perpendicular to this force?; what are the coordinates of the force in the rotated coordinate system $x' - y'$?; what are the coordinates of the force if its direction is rotated? These last two questions are related.

1.2.2 Linear Systems

Consider the following simple example. You probably saw something like this in high school.

Example 1.1 *Bob and Sue together have 12 dollars. Sue has 2 dollars more than Bob. How much money do each have?* You can probably guess the solution by trial and error, but let us proceed a bit more formally. Let x be the amount of

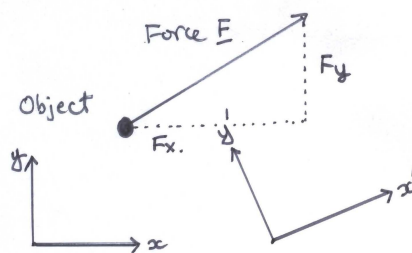


Figure 1.1: Force vector and coordinate systems.

money Sue has and y the amount Bob has. The two statements in the example can be written mathematically as

$$\begin{aligned}x + y &= 12 \\x - y &= 2.\end{aligned}$$

The equations above are a *linear system* for the unknowns x and y . A technique that can be used to solve the system (that is, determine the values of x and y that *simultaneously* solve both equations above) is substitution. The second equation can be written as

$$x = y + 2$$

This can be substituted into the first equation above, *eliminating* y from the problem

$$(y + 2) + y = 12 \text{ or } 2y + 2 = 12 \text{ so } 2y = 10 \text{ so } y = 5$$

The value of $y = 5$ determines $x = 7$ from either of the original relationships. Thus it is determined that Bob has 5 dollars and Sue has 7.

Often (but not always) a linear system of n equations for n unknowns has a unique solution. The example above was for the case $n = 2$. However, the substitution technique used above becomes impractical when n is larger than 3. In this course you will learn the Gaussian Elimination technique to solve linear systems. This systematic method can find all solutions when there are any and also determine if the system has no solutions. This method can be implemented in numerical software and used to solve very large systems.

1.2.3 Eigen-analysis

The final subject of the course is eigen-analysis of matrices and its applications. A simple, motivational example comes from the study of discrete dynamical systems. Consider a sequence of values

$$x_0, x_1, \dots, x_n \dots$$

where the index n is a time level. Suppose that x_n is determined by the previous value x_{n-1} in the same way for every n , that is

$$x_n = f(x_{n-1}) \quad \text{for every } n \geq 1 \quad (1.1)$$

for a given function f . This could describe the population number x_n of a species at year n . The simple model assumes that the population the next year only depends on the population this year through the function f . If the initial value x_0 were given, then the values $x_1, x_2 \cdots x_n \cdots$ could be determined using (1.1) repeatedly. A linear problem arises when we take the specific example $f(x) = ax$ where a is a given constant. In this case, it is easy to compute the entries of the sequence:

$$\begin{aligned} x_1 &= f(x_0) = ax_0 \\ x_2 &= f(x_1) = f(ax_0) = a^2x_0 \\ &\vdots \\ x_n &= f(x_{n-1}) = a^n x_0 \end{aligned}$$

For this example, we can determine how the sequence behaves very well because we have an expression for x_n above that is easy to understand. There are several cases:

1. If $x_0 = 0$ then $x_n = 0$ for all n .
2. if $|a| < 1$ then $\lim_{n \rightarrow \infty} x_n = 0$.
3. if $a = 1$ then $x_n = x_0$ for all n .
4. if $a = -1$ then the values alternate in sign: x_n has the value x_0 if n is even, $-x_0$ if n is odd.
5. if $|a| > 1$ and $x_0 \neq 0$ then the values of the sequence grow in absolute value as $n \rightarrow \infty$.

Linear discrete dynamical systems for vectors are also of interest. In these cases, multiplication by the number a in the example above is replaced by multiplication by a matrix \mathbf{A} . Eigen-analysis of the matrix \mathbf{A} allows one to understand how the system behaves as $n \rightarrow \infty$ in a similar way to the simple example above. Eigen-analysis naturally involves complex arithmetic, so this is introduced in the notes beforehand.

1.3 About The Computer Labs

The course includes six one-hour computer labs. These are given to small groups of students every other week starting in the *second* week of the term. It is a good idea to read through the lab notes before going to the lab so you are ready to begin in the lab. After your first lab, you will be able to go to the lab rooms

in open hours to improve your MATLAB skills and to prepare for upcoming labs. Computer lab material including your knowledge of MATLAB commands will be tested on midterms and the final exam.

There are two main goals for the labs. The first is to gain familiarity with the computational tool, MATLAB, that is commonly used in later courses and Engineering careers. The second is to be able to solve larger, more interesting applied problems that would otherwise be inaccessible using analytic methods. Seeing the algorithms of MATLAB in action may also help you understand the underlying mathematical concepts you see in the lectures.

1.4 About These Notes

2004: The first version of these notes was written by Richard Froese for Math 152 taught in the Spring of 2004. There are many text books on elementary linear algebra material, but none have the material in the order we want for Math 152 for Applied Science students. These notes stress geometric concepts in two and three dimensional space. They also treat applications and numerical approximation using MATLAB in more detail than most texts. For this reason, the authors have felt it was worthwhile to maintain and improve these notes for Math 152. Additionally, we believe it is a social benefit for students to have access to this material without having to purchase an expensive, commercial text.

2007: An update to the notes was made by Richard Froese for the course in 2007, including solutions to the exercises.

2009: The version written for the 2009 course had some updates by Brian Wetton: this introductory chapter, some additional comments on MATLAB commands, a reworked section on linear systems arising from electrical networks, and additional problems and solutions. In addition, the notes were converted to standard \LaTeX format to make them easier to maintain.

2010: Substantial revisions for the notes for 2010 were done by Ignacio Rozada, who was supported by a UBC Skylight grant over the Summer of 2009 to add the problems and solutions used in weekly assignments the previous year. He also added additional MATLAB material. Brian Wetton added additional notes to Chapter 4 on the use of matrix multiplication and inverses in the derivation of solutions to the “fundamental problem” of resistor networks.

2015: Brian Wetton added many additional examples to Chapter 2 and added more formal definitions of the key concepts: linear combination, span, linear independence, basis and dimension. He also added material in Chapter 5 on the use of complex arithmetic in complex linear systems and determinants. Joel Feldman contributed an additional topics section to Chapter 3 on the checksum technique. Hand drawn resistor network diagrams were

replaced throughout the text by Egor Dontsov. The material on complex numbers was split off into a separate chapter.

Further revisions are planned. There will be additional MATLAB material added to the notes and additional problems and solutions. The Chapter on complex numbers needs some revision. If you have any suggestions for additional material or ways to improve the presentation, please send your ideas to Brian Wetton (wetton@math.ubc.ca).

Chapter 2

Vectors and Geometry

2.1 Chapter Introduction

This chapter contains an introduction to vectors, which correspond to points in two, three and higher dimensional spaces. In this chapter, you will become familiar with basic vector operations such as addition, scalar multiplication, length, the dot product, and the cross product (for three dimensional vectors). Vector representation of lines in 2D and 3D and planes in 3D are presented. Criteria for when such objects intersect at unique points is given in terms of determinants. This geometric presentation motivates our study of these kind of problems in higher dimensional settings in later chapters. Throughout this chapter, MATLAB commands are introduced that perform the operations described in the text. For 2D and 3D problems, using MATLAB is only a convenience. For higher dimensions, doing the computations by hand (even with a calculator) is impractical, and a computational framework like MATLAB is essential to be able to solve these problems.

2.2 Vectors

Vectors are used to describe quantities that have both a magnitude and a direction. You are probably familiar with vector quantities in two and three dimensions, such as forces and velocities.

Later in this course we will see that vectors can also describe the configuration of a mechanical system of weights and springs, or the collections of voltages and currents in an electrical circuit. These more abstract vector quantities are not so easily visualized since they take values in higher dimensional spaces.

We begin this course by discussing the geometry of vectors in two and three dimensions. In two and three dimensions, vectors can be visualized as arrows. Before we can draw a vector, we have to decide where to place the tail of the vector. If we are drawing forces, we usually put the tail of the vector at the

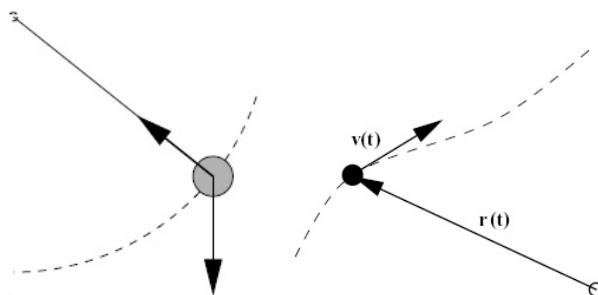


Figure 2.1: Forces acting on a pendulum (left) and position and velocity of a particle (right)

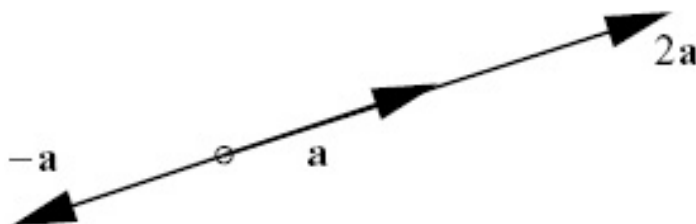


Figure 2.2: Scalar multiplication.

place where the force is applied. For example, in Figure 2.1 (left) the forces acting on a pendulum bob are gravity and the restraining force along the shaft.

If we are drawing the velocity of a particle at a given time, we would place the tail of the velocity vector $\mathbf{v}(t)$ at the position of the particle at that time as shown in Figure 2.1 (right). Once we have chosen a starting point for the tails of our vectors (i.e., an origin for space), every point in space corresponds to exactly one vector, namely the vector whose tail is at the origin and whose head is at the given point. For example, in Figure 2.1 (right) we have chosen an arbitrary point as the origin (marked with a circle) and identified the position of the particle with the vector $\mathbf{r}(t)$.

2.2.1 Multiplication by a number and vector addition

There are two basic operations defined for vectors. One is multiplication of a vector by a number (also called scalar multiplication). The other is addition of two vectors.

A vector \mathbf{a} can be multiplied by a number (or scalar) s to produce a new vector $s\mathbf{a}$. If s is positive then $s\mathbf{a}$ points in the same direction as \mathbf{a} and has length s times the length of \mathbf{a} . This is shown in Figure 2.2. If s is negative then $s\mathbf{a}$ points in the direction opposite to \mathbf{a} and has length $|s|$ times the length of \mathbf{a} .



Figure 2.3: Vector Addition.

To add two vectors \mathbf{a} and \mathbf{b} and we draw the parallelogram that has \mathbf{a} and \mathbf{b} as two of its sides as shown in Figure 2.3. The vector $\mathbf{a} + \mathbf{b}$ has its tail at the origin and its head at the vertex of the parallelogram opposite the origin. Alternatively we can imagine sliding (or translating) one of the vectors, without changing its direction, so that its tail sits on the head of the other vector. (In the diagram we translated the vector \mathbf{a} .) The sum $\mathbf{a} + \mathbf{b}$ is then the vector whose tail is at the origin and whose head coincides with the vector we moved.

Example 2.1 Describe and sketch the following set of points $\{s\mathbf{a} : s \in \mathbb{R}\}$ (that is, the set of all scalar multiples of \mathbf{a}) where \mathbf{a} is a non-zero vector in \mathbb{R}^2 . The set is a straight line going through the origin with direction \mathbf{a} as shown in Figure 2.4.

2.2.2 Co-ordinates

In order to do calculations with vectors, we have to introduce co-ordinate axes. Once we have chosen in what directions the co-ordinate axes will lie, we can specify a vector by giving its components in the co-ordinate directions.

In the Figure 2.5 we see two choices of x and y axes. For the first choice of axes, the vector \mathbf{a} has co-ordinates $[5, 3]$ and for the second choice of axes the co-ordinates are $[\sqrt{34}, 0]$. In a given problem, it makes sense to choose the axes so that at least some of the vectors have a simple representation. For example, in analyzing the forces acting on a pendulum, we would either choose the y axis either to be vertical, or to lie along the shaft of the pendulum.

We can choose to write the co-ordinates of a vector in a row, as above, or in a column, like

$$\begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

Later on, we will almost always write vectors as columns. But in this chapter we will write vectors as rows. Writing vectors as rows saves space on the page but we will learn later how to write vectors in row form even when we want them to be column vectors for other reasons. **Note:** When writing vector coordinates by

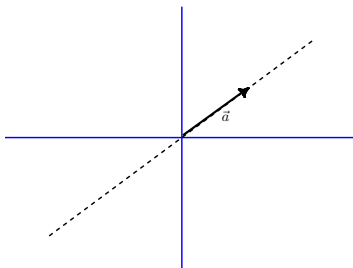


Figure 2.4: Figure for Example 2.1.

hand or in this text, either square or round brackets can be used. However, when using MATLAB, vectors must be created with square brackets (round brackets are used for other purposes). Other packages such as the online homework system, WeBWorK, may require different syntax.

A convenient way to choose the co-ordinate axes is to specify unit vectors (that is, vectors of length one) that lie along each of the axes. These vectors are called standard basis vectors and are denoted \mathbf{i} and \mathbf{j} (in two dimensions) and \mathbf{i} , \mathbf{j} and \mathbf{k} (in three dimensions). These vectors are shown in Figure 2.6. Sometimes they are also denoted \mathbf{e}_1 and \mathbf{e}_2 or \hat{i} and \hat{j} (in two dimensions) and \mathbf{e}_1 , \mathbf{e}_2 and \mathbf{e}_3 or \hat{i} , \hat{j} and \hat{k} (in three dimensions). Different application areas have different conventions.

The unit vectors have co-ordinates

$$\begin{aligned}\mathbf{i} &= \mathbf{e}_1 = [1, 0] \\ \mathbf{j} &= \mathbf{e}_2 = [0, 1]\end{aligned}$$

in two dimensions, and

$$\begin{aligned}\mathbf{i} &= \mathbf{e}_1 = [1, 0, 0] \\ \mathbf{j} &= \mathbf{e}_2 = [0, 1, 0] \\ \mathbf{k} &= \mathbf{e}_3 = [0, 0, 1]\end{aligned}$$

in three dimensions.

Often, we make no distinction between a vector and its co-ordinate representation. In other words, we regard the co-ordinate axes as being fixed once

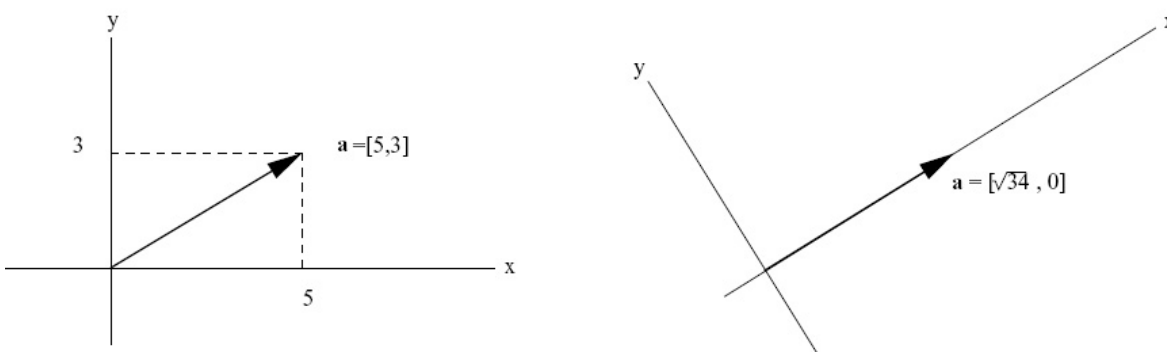


Figure 2.5: Two choices of co-ordinate axes.

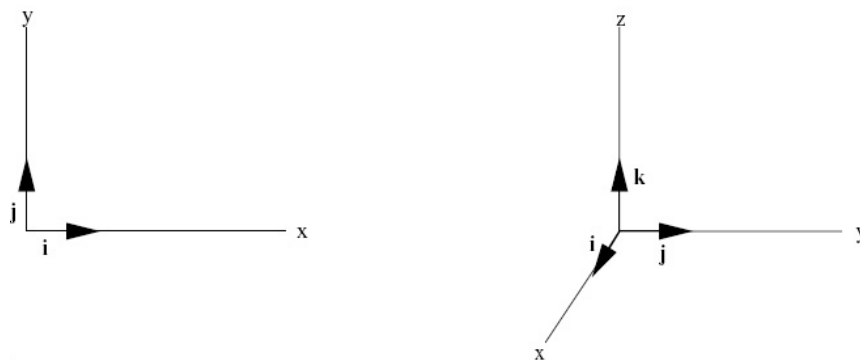


Figure 2.6: Unit vectors in 2D (left) and 3D (right).

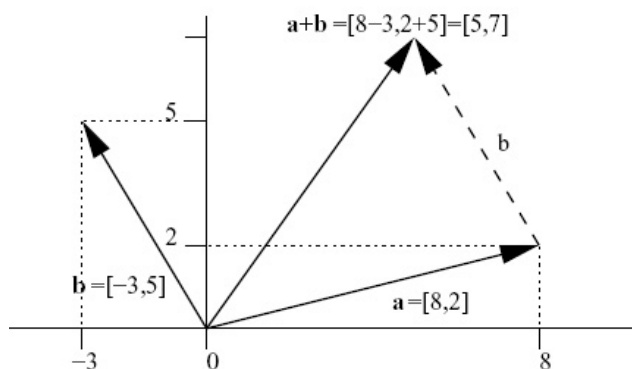


Figure 2.7: Adding vector co-ordinates.

and for all. Then a vector in two dimensions is simply a list of two numbers (the components) $[a_1, a_2]$, and a vector in three dimensions is a list of three numbers $[a_1, a_2, a_3]$. Vectors in higher dimensions are now easy to define. A vector in n dimensions is a list of n numbers $[a_1, a_2, \dots, a_n]$.

When a vector is multiplied by a number, each component is scaled by the same amount. Thus if $\mathbf{a} = [a_1, a_2]$, then

$$\begin{aligned} s\mathbf{a} &= s[a_1, a_2] \\ &= [sa_1, sa_2] \end{aligned}$$

Similarly, when two vectors are added, their co-ordinates are added component-wise. So if $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$, then

$$\begin{aligned} \mathbf{a} + \mathbf{b} &= [a_1, a_2] + [b_1, b_2] \\ &= [a_1 + b_1, a_2 + b_2] \end{aligned}$$

This is shown in Figure 2.7.

The analogous formulae hold in three (and higher dimensions). If $\mathbf{a} = [a_1, a_2, \dots, a_n]$ and $\mathbf{b} = [b_1, b_2, \dots, b_n]$, then

$$\begin{aligned} s\mathbf{a} &= s[a_1, a_2, \dots, a_n] \\ &= [sa_1, sa_2, \dots, sa_n] \\ \mathbf{a} + \mathbf{b} &= [a_1, a_2, \dots, a_n] + [b_1, b_2, \dots, b_n] \\ &= [a_1 + b_1, a_2 + b_2, \dots, a_n + b_n] \end{aligned}$$

Example 2.2 Sketch axes x_1 - x_2 . Add the vectors $(1,1)$ and $(2,-1)$ to your sketch. Draw these vectors with base point at the origin. Now add the vector $(1,-2)$ to your sketch, starting at the base point $(1,1)$. That is, draw the vector with components 1 to the right and 2 down starting at $(1,1)$. **Note:** your sketch should show graphically that $(1,1)+(1,-2)=(2,-1)$. See Figure 2.8.

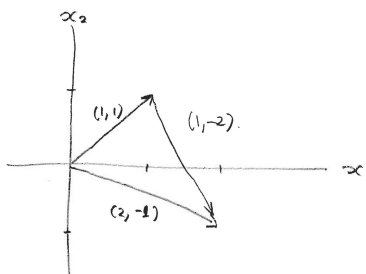


Figure 2.8: Figure for example 2.2

2.2.3 Properties of vector addition and scalar multiplication

Let $\mathbf{0}$ denote the zero vector. This is the vector all of whose components are zero. The following properties are intuitive and easy to verify.

1. $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
2. $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$
3. $\mathbf{a} + \mathbf{0} = \mathbf{a}$
4. $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$
5. $s(\mathbf{a} + \mathbf{b}) = (s\mathbf{a} + s\mathbf{b})$
6. $(s + t)\mathbf{a} = s\mathbf{a} + t\mathbf{a}$
7. $(st)\mathbf{a} = s(t\mathbf{a})$
8. $1\mathbf{a} = \mathbf{a}$

They follow from similar properties which hold for numbers. For example, for numbers a_1 and b_1 we know that $a_1 + b_1 = b_1 + a_1$. Thus

$$\begin{aligned} \mathbf{a} + \mathbf{b} &= [a_1, a_2] + [b_1, b_2] \\ &= [a_1 + b_1, a_2 + b_2] = [b_1 + a_1, b_2 + a_2] \\ &= [b_1, b_2] + [a_1, a_2] = \mathbf{b} + \mathbf{a}, \end{aligned}$$

so property 1 holds. Convince yourself that the rest of these properties are true. (What is the vector $-\mathbf{a}$?). It might seem like a waste of time fussing over obvious properties such as these. However, we will see when we come to the cross product and matrix product, that sometimes such “obvious” properties turn out to be false! It is important to know what the allowable operations for vectors (and matrices which we will see in later chapters) are since we will be doing algebra to solve matrix and vector equations. We will take special care to highlight the operations that have different properties from the scalar case.

2.2.4 MATLAB: basic scalar and vector operations

There are MATLAB computer labs that accompany this course at UBC. You will be given an account on the Mathematics Department undergraduate network and instructions on how to access the computers in the labs and start the MATLAB application. A freely available google application, Octave, can execute all of the commands in the course with the same syntax. In the command window at the prompt `>>` (MATLAB) or `▷` (Octave), you can type MATLAB commands directly. Some basic commands are given below

assignment: Scalar and vector variables can be assigned using the “=” operator. For example

```
a = 2
```

followed by `<enter>` assigns the scalar value of 2 to the variable `a`. The result of the command is printed out although this can be suppressed by using a colon at the end of the command:

```
a = 2;
```

Here `a` is still assigned the value of 2 but no output is generated. Vector variables are assigned with the following notation:

```
b = [1 2];
```

Note that `b = [1, 2]` has the same meaning in MATLAB, i.e., numbers separated by a comma or a space imply row vectors. For column vectors, the entries have to be separated by semicolons:

```
b1 = [2; 3];
```

Note also that there are no special distinctions between the names of scalar and vector variables.

addition: Both scalar and vector addition can be done with the “+” operator. Keeping the values of scalar `a` and vector `b` above, we enter the commands

```
a2 = 5;  
b2 = [2 9];  
a+a2  
c = b+b2;
```

The first two lines above assign a new scalar and vector. The third line prints out the answer 7 (2+5). The last line assigns the resulting vector [3 11] ($[1\ 2] + [2\ 9]$) to the new vector `c` but prints nothing.

scalar multiplication Scalar multiplication (of vectors and other scalars) is implemented using the “*” command. Using the variables defined above,

```
a*a2
a*b
```

would result in 10 (2 times 5) and [2 4] (2 times [1 2]). The “*” command also implements matrix-vector and matrix-matrix multiplication discussed later in the course. Vector-vector multiplication (dot products and cross products) are implemented using different commands as discussed in the next section.

other commands: There are many useful functions built in to MATLAB such as `sqrt` (square root), `cos` (cosine, taking an argument in radians), `acos` (inverse cosine, giving an result in radians) and many more. They are called as follows

```
sqrt(2)
```

which will return $\sqrt{2}$ to 4 decimal places. Type `help` followed by a command name gives you a description of that command. Try typing `help atan2` since `atan2` is a pretty useful function. These MATLAB functions can take vector arguments, they act on each entry of the vector. For example

```
sqrt([1 4])
```

will produce the vector [1 2].

2.2.5 Problems

Problem 2.1 Sketch axes x_1 - x_2 . Add the vectors $(2,2)$ and $(1,-1)$ to your sketch. Draw these vectors with base point at the origin. Now add the vector $(1,-1)$ to your sketch, starting at the base point $(2,2)$. That is, draw the vector with components 1 to the right and 1 down starting at $(2,2)$. **Note:** your sketch should show graphically that $(2,2)+(1,-1)=(3,1)$.

Problem 2.2 Let \mathbf{a} , \mathbf{b} and \mathbf{c} be fixed non-zero vectors. Describe and sketch the following sets of points in two and three dimensions:

- (i) $\{\mathbf{sa} : s \in \mathbf{R}\}$ (i.e., the set of all scalar multiples of \mathbf{a})
- (ii) $\{\mathbf{sa} : s > 0\}$ (i.e., the set of all positive scalar multiples of \mathbf{a})
- (iii) $\{\mathbf{b} + \mathbf{sa} : s \in \mathbf{R}\}$
- (iv) $\{\mathbf{sa} + \mathbf{tb} : s, t \in \mathbf{R}\}$
- (v) $\{\mathbf{c} + \mathbf{sa} + \mathbf{tb} : s, t \in \mathbf{R}\}$

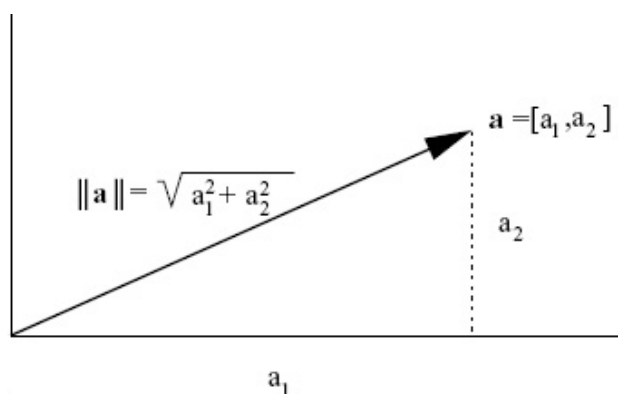


Figure 2.9: Pythagorean Formula.

Problem 2.3 Describe the vectors $\mathbf{a} - \mathbf{b}$ and $\mathbf{b} - \mathbf{a}$.

Problem 2.4 Find an expression for the midpoint between \mathbf{a} and \mathbf{b} . Find an expression for a point one third of the way between \mathbf{a} and \mathbf{b} .

Problem 2.5 Find an expression for the line segment joining \mathbf{a} and \mathbf{b} .

2.3 Geometrical Aspects of Vectors

2.3.1 Length of a vector

It follows from the Pythagorean formula that the length $\|\mathbf{a}\|$ of $\mathbf{a} = [a_1, a_2]$ satisfies $\|\mathbf{a}\|^2 = a_1^2 + a_2^2$. This is shown in Figure 2.9.

Thus

$$\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2}.$$

Similarly, for a vector $\mathbf{a} = [a_1, a_2, a_3]$ in three dimensions,

$$\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

The distance between two vectors \mathbf{a} and \mathbf{b} is the length of the difference $\mathbf{b} - \mathbf{a}$.

2.3.2 The dot product

The dot product of two vectors is defined in both two and three dimensions (actually in any dimension). The result is a number. Two main uses of the dot product are testing for orthogonality and computing projections.

The dot product of $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$ is given by

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2.$$

Similarly, the dot product of $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$ is given by

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

The properties of the dot product are as follows:

0. If \mathbf{a} and \mathbf{b} are vectors, then $\mathbf{a} \cdot \mathbf{b}$ is a number.
1. $\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2$.
2. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$.
3. $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$.
4. $s(\mathbf{a} \cdot \mathbf{b}) = (s\mathbf{a}) \cdot \mathbf{b}$.
5. $\mathbf{0} \cdot \mathbf{a} = 0$.
6. $\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \cos(\theta)$, where θ is angle between \mathbf{a} and \mathbf{b} .
7. $\mathbf{a} \cdot \mathbf{b} = 0 \iff \mathbf{a} = \mathbf{0}$ or $\mathbf{b} = \mathbf{0}$ or \mathbf{a} and \mathbf{b} are orthogonal (i.e., perpendicular).

Properties 0 to 5 are easy consequences of the definitions. For example, to verify property 5 we write

$$\mathbf{0} \cdot \mathbf{a} = [0, 0, 0] \cdot [a_1, a_2, a_3] = 0a_1 + 0a_2 + 0a_3 = 0.$$

Property 6 is the most important property and is often taken as the definition of the angle θ between vectors \mathbf{a} and \mathbf{b} . Notice that our definition is given in terms of the components of the vectors, which depend on how we chose the co-ordinate axes. It is not at all clear that if we change the co-ordinate axis, and hence the co-ordinates of the vectors, that we will get the same answer for the dot product. However, property 6 says that the dot product only depends on the lengths of the vectors and the angle between them. These quantities are independent of how co-ordinate axes are chosen, and hence so is the dot product.

To show that property 6 holds we compute $\|\mathbf{a} - \mathbf{b}\|^2$ in two different ways. First of all, using properties 1 to 5, we have

$$\begin{aligned} \|\mathbf{a} - \mathbf{b}\|^2 &= (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) \\ &= \mathbf{a} \cdot \mathbf{a} - \mathbf{a} \cdot \mathbf{b} - \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} \\ &= \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\mathbf{a} \cdot \mathbf{b} \end{aligned} \tag{2.1}$$

(Which properties were used in each step?) Next we compute $\|\mathbf{a} - \mathbf{b}\|$ as depicted in Figure 2.10 (left).

We mark the lengths of each of the line segments in Figure 2.10 (right). Using Pythagoras' theorem for the right angled triangle on the right of this diagram, we see that

$$\|\mathbf{a} - \mathbf{b}\|^2 = (\|\mathbf{a}\| - \|\mathbf{b}\| \cos(\theta))^2 + \|\mathbf{b}\|^2 \sin^2(\theta).$$

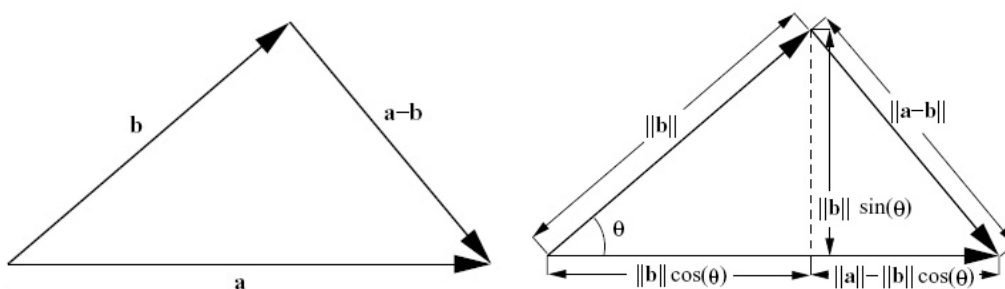


Figure 2.10: The vectors \mathbf{a} , \mathbf{b} and $\mathbf{a} - \mathbf{b}$ (left) Lengths of Segments (right).

Thus, using $\cos^2(\theta) + \sin^2(\theta) = 1$,

$$\begin{aligned} \|\mathbf{a} - \mathbf{b}\|^2 &= \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 \cos^2(\theta) - 2\|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta) + \|\mathbf{b}\|^2 \sin^2(\theta) \\ &= \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta) \end{aligned} \quad (2.2)$$

Actually, this is just the cosine law applied to the triangle in Figure 2.10 and you may have been able to write (2.2) directly. Now we equate the two expressions (2.1, 2.2) for $\|\mathbf{a} - \mathbf{b}\|^2$. This gives

$$\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta)$$

Subtracting $\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2$ from both sides and dividing by -2 now yields

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta).$$

This proves property 6.

Property 7 now follows directly from 6. If $\mathbf{a} \cdot \mathbf{b} = 0$ then $\|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta) = 0$ so either $\|\mathbf{a}\| = 0$, in which case $\mathbf{a} = \mathbf{0}$, or $\|\mathbf{b}\| = 0$, in which case $\mathbf{b} = \mathbf{0}$, or $\cos(\theta) = 0$, which implies that $\theta = \pi/2$ (since θ lies between 0 and π). This implies \mathbf{a} and \mathbf{b} are orthogonal.

Property 6 can be used to compute the angle between two vectors as shown in the example below.

Example 2.3 *What is the angle between the vectors whose tails lie at the centre of a cube and whose heads lie on adjacent vertices?* To compute this take a cube of side length 2 and centre it at the origin, so that the vertices lie at the points $[\pm 1, \pm 1, \pm 1]$. Then we must find the angle between $\mathbf{a} = [1, 1, 1]$ and $\mathbf{b} = [-1, 1, 1]$. Since

$$\mathbf{a} \cdot \mathbf{b} = -1 + 1 + 1 = 1 = \|\mathbf{a}\|\|\mathbf{b}\| \cos(\theta) = \sqrt{3}\sqrt{3} \cos(\theta)$$

we obtain

$$\theta = \arccos(1/3) \sim 1.231 \quad (\sim 70.5^\circ)$$

Example 2.4 For what value (or values) of s is the vector $[1, 1, s]$ perpendicular to $[1, 5, 3]$? The dot product of the two vectors must be zero for them to be orthogonal. We compute

$$[1, 1, s] \cdot [1, 5, 3] = 1 + 5 + 3s$$

so $6 + 3s = 0$ for orthogonality, $s = -2$.

Here is an example to review the basic operations on vectors we know so far.

Example 2.5 Consider the vectors $\mathbf{a} = (2, 3)$ and $\mathbf{b} = (1, -3)$ in \mathbb{R}^2 . Compute the following:

- (a) $\mathbf{a} + \mathbf{b}$
- (b) $3\mathbf{a}$
- (c) $2\mathbf{a} + 4\mathbf{b}$
- (d) $\mathbf{a} \cdot \mathbf{b}$
- (e) $\|\mathbf{b}\|$

Solutions:

- (a) $\mathbf{a} + \mathbf{b} = (2, 3) + (1, -3) = (3, 0)$
- (b) $3\mathbf{a} = 3(2, 3) = (6, 9)$
- (c) $2\mathbf{a} + 4\mathbf{b} = 2(2, 3) + 4(-1, 3) = (4, 6) + (4, -12) = (8, -6)$
- (d) $\mathbf{a} \cdot \mathbf{b} = (2, 3) \cdot (1, -3) = 2 - 9 = -7$.
- (e) $\|\mathbf{b}\| = \sqrt{1^2 + (-3)^2} = \sqrt{10}$.

2.3.3 Projections and Unit Vectors

Suppose \mathbf{a} and \mathbf{b} are two vectors. The projection of \mathbf{a} in the direction of \mathbf{b} , denoted $\text{proj}_{\mathbf{b}}\mathbf{a}$, is the vector in the direction of \mathbf{b} whose length is determined by drawing a line perpendicular to \mathbf{b} that goes through \mathbf{a} . In other words, the length of $\text{proj}_{\mathbf{b}}\mathbf{a}$ is the component of \mathbf{a} in the direction of \mathbf{b} . This is shown in Figure 2.11

To compute $\text{proj}_{\mathbf{b}}\mathbf{a}$, we first note that it is a multiple of \mathbf{b} . Thus $\text{proj}_{\mathbf{b}}\mathbf{a} = s\mathbf{b}$ for some number s . To compute s , we use the fact that the vector $\text{proj}_{\mathbf{b}}\mathbf{a} - \mathbf{a}$ (along the dotted line in the diagram) is orthogonal to \mathbf{b} . Thus $(\text{proj}_{\mathbf{b}}\mathbf{a} - \mathbf{a}) \cdot \mathbf{b} = 0$, or $(s\mathbf{b} - \mathbf{a}) \cdot \mathbf{b} = 0$, or $s = \mathbf{a} \cdot \mathbf{b} / \mathbf{b} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{b} / \|\mathbf{b}\|^2$. Thus

$$\text{proj}_{\mathbf{b}}\mathbf{a} = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|^2} \mathbf{b}. \quad (2.3)$$

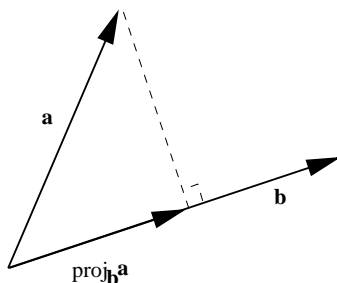


Figure 2.11: Projection.

If \mathbf{b} is a unit vector (i.e., $\|\mathbf{b}\| = 1$, length 1) this expression is even simpler. We will use the notation $\hat{\mathbf{b}}$ to denote unit vectors. In this case

$$\text{proj}_{\hat{\mathbf{b}}}\mathbf{a} = (\mathbf{a} \cdot \hat{\mathbf{b}})\hat{\mathbf{b}}.$$

It is easy to find a unit vector $\hat{\mathbf{b}}$ that points in the same direction as any nonzero vector \mathbf{b} , simply compute

$$\hat{\mathbf{b}} = \frac{1}{\|\mathbf{b}\|}\mathbf{b}.$$

Since $\hat{\mathbf{b}}$ is a (positive) scalar multiple of \mathbf{b} in the above formula, it clearly points in the same direction as \mathbf{b} . A straight forward calculation shows that $\hat{\mathbf{b}}$ has unit length (try it!).

Example 2.6 Find the unit vector $\hat{\mathbf{b}}$ that points in the same direction as $\mathbf{b} = [1, 2]$. We compute

$$\|\mathbf{b}\| = \sqrt{1^2 + 2^2} = \sqrt{5}$$

and then

$$\hat{\mathbf{b}} = \frac{1}{\sqrt{5}}[1, 2] = [1/\sqrt{5}, 2/\sqrt{5}].$$

Projections are useful for computing the components of a vector in various directions. An easy example is given by the co-ordinates of a vector. These are simply the components of a vector in the direction of the standard basis vectors. So in two dimensions

$$\begin{aligned} a_1 &= \mathbf{a} \cdot \mathbf{i} = [a_1, a_2] \cdot [1, 0] \\ a_2 &= \mathbf{a} \cdot \mathbf{j} = [a_1, a_2] \cdot [0, 1] \end{aligned}$$

Example 2.7 Let $\mathbf{a} = [1, 0, 2]$ and $\mathbf{b} = [0, 5, 2]$. Compute $\text{proj}_{\mathbf{b}}\mathbf{a}$ (the projection of \mathbf{a} in the direction of \mathbf{b}). We use (2.3) and first compute

$$\|\mathbf{b}\|^2 = 5^2 + 2^2 = 29 \quad \text{and} \quad \mathbf{a} \cdot \mathbf{b} = 4,$$

then

$$\text{proj}_{\mathbf{b}}\mathbf{a} = \frac{4}{29}[0, 5, 2] = [0, 20/29, 8/29].$$

2.3.4 MATLAB: norm and dot commands

MATLAB has built-in functions that implement most of the mathematical operations introduced this course. For example, the commands

`norm(a)` returns the length (norm) of the vector \mathbf{a} .

`dot(a,b)` returns the dot product of the vectors \mathbf{a} and \mathbf{b} (if the vectors do not have the same length, an error results as you would expect).

Using these commands and scalar multiplication of vectors, a projection of \mathbf{a} onto the direction \mathbf{b} can be implemented:

`(dot(a,b)/norm(b)^2)*b`

where `/` denotes division (of scalar quantities in this case) and `^ p` gives the p 'th power of a quantity.

2.3.5 Problems

Problem 2.6 Consider the vectors $\mathbf{a} = (1, 2)$ and $\mathbf{b} = (1, -2)$ in \mathbb{R}^2 (the set of vectors with 2 components). Compute the following:

1. $\mathbf{a} + \mathbf{b}$
2. $2\mathbf{a}$
3. $\mathbf{a} - \mathbf{b}$
4. $\mathbf{a} \cdot \mathbf{b}$
5. $\|\mathbf{b}\|$

Problem 2.7 A circle in the x_1 - x_2 plane has centre at $(2, 5)$. A given point on its circumference is $(3, 3)$. Write an equation that describes all the points (x_1, x_2) on the circle.

Problem 2.8 Find the equation of a sphere centred at $\mathbf{a} = [a_1, a_2, a_3]$ with radius r . (Hint: the sphere is the set of points $\mathbf{x} = [x_1, x_2, x_3]$ whose distance from \mathbf{a} is r)

Problem 2.9 Find the equation of a sphere if one of its diameters has endpoints $[2, 1, 4]$ and $[4, 3, 10]$

Problem 2.10 Compute the dot product of the vectors \mathbf{a} and \mathbf{b} and find the angle between them.

(i) $\mathbf{a} = [1, 2]$, $\mathbf{b} = [-2, 3]$

(ii) $\mathbf{a} = [-1, 2]$, $\mathbf{b} = [1, 1]$

(iii) $\mathbf{a} = [1, 1]$, $\mathbf{b} = [2, 2]$

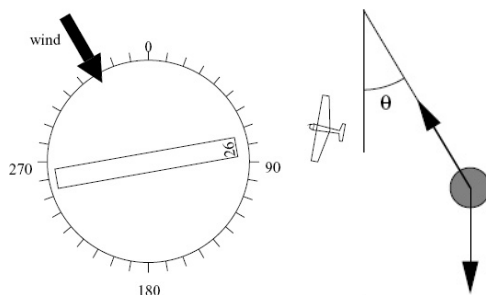


Figure 2.12: Runway diagram for problem 2.16 (left) and the pendulum of problem 2.17 (right)

(iv) $\mathbf{a} = [1, 2, 1]$, $\mathbf{b} = [-1, 1, 1]$

(v) $\mathbf{a} = [-1, 2, 3]$, $\mathbf{b} = [3, 0, 1]$

Problem 2.11 Let $\mathbf{a} = (1, 1, 1)$ and $\mathbf{b} = (3, 1, -2)$. Compute the following:

1. The angle between \mathbf{a} and \mathbf{b} .
2. $\text{proj}_{\mathbf{a}}\mathbf{b}$ (the projection of \mathbf{b} in the direction of \mathbf{a}).

Problem 2.12 Let $\mathbf{a} = (1, 4, 0)$ and $\mathbf{b} = (2, -1, 5)$. Compute the following:

- (a) The angle between \mathbf{a} and \mathbf{b} .
- (b) $\text{proj}_{\mathbf{a}}\mathbf{b}$ (the projection of \mathbf{b} in the direction of \mathbf{a}).

Problem 2.13 For which value of s is the vector $[1, 2, s]$ orthogonal to $[-1, 1, 1]$?

Problem 2.14 Does the triangle with vertices $[1, 2, 3]$, $[4, 0, 5]$ and $[3, 4, 6]$ have a right angle?

Problem 2.15 Determine the values of c_1 and c_2 such that the vector $[c_1 \ 1 \ c_2]$ is a scalar multiple of $[2 \ -2 \ 3]$.

Problem 2.16 An air-plane with an approach speed of 70 knots is on approach to runway 26 (i.e., pointing in the direction of 260 degrees). This is shown in Figure 2.12 (left). If the wind is from 330 degrees at 10 knots, what heading should the pilot maintain to stay lined up with the runway? What is the groundspeed of the air-plane?

Problem 2.17 Suppose the angle of the pendulum shaft makes an angle of θ with the vertical direction as shown in Figure 2.12 (right). The force of gravity has magnitude (length) equal to mg and points downwards. The force along the shaft of the pendulum acts to keep the shaft rigid, i.e., the component of the total force along the shaft is zero. Write down the co-ordinates of the two forces and the total force using two different sets of co-ordinate axes — one horizontal and vertical, and one parallel to and orthogonal to the shaft of the pendulum.

Problem 2.18 (*Matlab*) In Matlab code, if one defines a vector $\mathbf{a} = [a_1, a_2, \dots, a_n]$, with the a_i 's being any numbers, the output of typing $\mathbf{a}(j)$ would be a_j .

Suppose that you have a two element vector. How would you write a line of Matlab code to compute the norm of the vector, without using the `norm` or `dot` commands?

2.4 Determinants and the Cross Product

2.4.1 The determinant in two and three dimensions

The determinant is a number that is associated with a square matrix, that is, a square array of numbers.

In two dimensions it is defined by

$$\det \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} = a_1 b_2 - a_2 b_1.$$

Example 2.8 Find the determinant of

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

Using the formula above, we have that the determinant is

$$1 \times 4 - 2 \times 3 = -2.$$

The definition in three dimensions is

$$\begin{aligned} \det \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} &= a_1 \det \begin{bmatrix} b_2 & b_3 \\ c_2 & c_3 \end{bmatrix} - a_2 \det \begin{bmatrix} b_1 & b_3 \\ c_1 & c_3 \end{bmatrix} + a_3 \det \begin{bmatrix} b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} \\ &= a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1 \end{aligned}$$

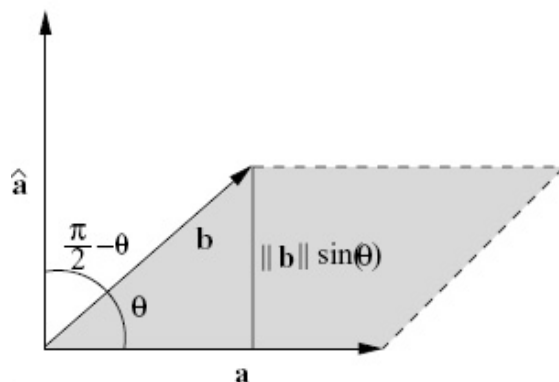
We want to determine the relationship between the determinant and the vectors $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$ (in two dimensions) and $\mathbf{a} = [a_1, a_2, a_3]$, $\mathbf{b} = [b_1, b_2, b_3]$ and $\mathbf{c} = [c_1, c_2, c_3]$ (in three dimensions). We will do the two dimensional case now, but postpone the three dimensional case until after we have discussed the cross product.

Example 2.9 Find the determinant of

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

It is easiest to remember the formula using the first line in the equation above using the determinants of the corresponding 2×2 blocks of the matrix.

$$\det A = 1(2 \times 3 - 0) - 1(4 \times 3 - 2) + 0 = -4.$$

Figure 2.13: The vector $\hat{\mathbf{a}}$.

So let $\mathbf{a} = [a_1, a_2]$ and $\mathbf{b} = [b_1, b_2]$ be two vectors in the plane. Define

$$\mathbf{a}^\perp = [-a_2, a_1].$$

Notice that \mathbf{a}^\perp has the same length as \mathbf{a} , and is perpendicular to \mathbf{a} , since

$$\mathbf{a}^\perp \cdot \mathbf{a} = -a_2 a_1 + a_1 a_2 = 0.$$

There are exactly two vectors with these properties. The vector \mathbf{a}^\perp is the one that is obtained from \mathbf{a} by a counterclockwise rotation of $\pi/2$ (i.e., 90°). To see this, notice that if \mathbf{a} lies in the first quadrant (that is, $a_1 > 0$ and $a_2 > 0$) then \mathbf{a}^\perp lies in the second quadrant, and so on. Later in the course we will study rotations and this will be a special case. Notice now that the determinant can be written as a dot product.

$$\mathbf{a}^\perp \cdot \mathbf{b} = -a_2 b_1 + a_1 b_2 = \det \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$$

We want to use the geometric formula for the dot product of \mathbf{a}^\perp and \mathbf{b} . Let θ be the angle between \mathbf{a} and \mathbf{b} and $\pi/2 - \theta$ be the angle between \mathbf{a}^\perp and \mathbf{b} , as shown in Figure 2.13.

Using the geometric meaning of the dot product, we obtain

$$\begin{aligned} \det \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} &= \mathbf{a}^\perp \cdot \mathbf{b} \\ &= \|\hat{\mathbf{a}}\| \|\mathbf{b}\| \cos(\pi/2 - \theta) \\ &= \|\mathbf{a}\| \|\mathbf{b}\| \sin(\theta) \end{aligned}$$

We need to be a bit careful here. When we were discussing the dot product, we always assumed that the angle between two vectors was in the range 0 to π . In fact, the geometric formula for the dot product is not sensitive to how we

measure the angle. Suppose that instead of θ in the range 0 to π we use $\theta_1 = -\theta$ (measuring the angle “backwards”) or $\theta_2 = 2\pi - \theta$ (measuring the angle going the long way around the circle). Since $\cos(\theta) = \cos(-\theta) = \cos(2\pi - \theta)$ we have

$$\mathbf{c} \cdot \mathbf{d} = \|\mathbf{c}\| \|\mathbf{d}\| \cos(\theta) = \|\mathbf{c}\| \|\mathbf{d}\| \cos(\theta_1) = \|\mathbf{c}\| \|\mathbf{d}\| \cos(\theta_2).$$

In other words, the geometric formula for the dot product still is true.

In the diagram above, we want to let the angle θ between \mathbf{a} and \mathbf{b} range between $-\pi$ and π . In this case the angle $\pi/2 - \theta$ between $\hat{\mathbf{a}}$ and \mathbf{b} is sometimes not in the range between 0 or 2π . But if this happens, then it is still the angle between \mathbf{a}^\perp and \mathbf{b} , just “backwards” or “the long way around.” Thus the geometric formula above still is correct.

Values of θ between 0 and π correspond to the situation where the direction of \mathbf{b} is obtained from the direction of \mathbf{a} by a counterclockwise rotation of less than π . This is the case in the diagram. On the other hand, θ between $-\pi$ and 0 corresponds to the case where a clockwise rotation of less than π is needed to get from the direction of \mathbf{a} to the direction of \mathbf{b} .

The quantity $\sin(\theta)$ can be positive or negative, depending on the orientations of \mathbf{a} and \mathbf{b} , but in any case the positive quantity $\|\mathbf{b}\| |\sin(\theta)|$ is the height of the parallelogram spanned by \mathbf{a} and \mathbf{b} if we take \mathbf{a} to be the base. In this case, the length of the base is $\|\mathbf{a}\|$. Recall that the area of a parallelogram is the length of the base times the height. Thus

$$\left| \det \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \right| = \text{Area of parallelogram spanned by } \mathbf{a} \text{ and } \mathbf{b}$$

The determinant is positive if $\sin(\theta)$ is positive, that is, if θ is positive. This is the case if the direction of \mathbf{b} is obtained by a counterclockwise rotation of half a circle or less from the direction of \mathbf{a} . Otherwise the determinant is negative.

Notice that the determinant whose rows are the components of two non-zero vectors \mathbf{a} and \mathbf{b} is zero exactly when the vectors \mathbf{a} and \mathbf{b} are pointing in the same direction, or in the opposite direction, that is, if one is obtained from the other by scalar multiplication. The sign of the determinant gives information about their relative orientation.

Example 2.10 Find the area \mathcal{A} of the parallelogram spanned by $\mathbf{a} = [1, 1]$ and $\mathbf{b} = [1, 3]$. Using the formula above, we know that

$$\mathcal{A} = \left| \det \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \right| = |3 - 1| = |2| = 2.$$

In addition, since the determinant is positive, we know that \mathbf{b} is counterclockwise to \mathbf{a} as can be seen graphically.

2.4.2 The cross product

Unlike the dot product, the cross product is only defined for vectors in three dimensions. And unlike the dot product, the cross product of two vectors is

another vector, not a number. If $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$, then $\mathbf{a} \times \mathbf{b}$ is a vector given by

$$\mathbf{a} \times \mathbf{b} = [a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1].$$

An easy way to remember this is to write down a 3×3 matrix whose first row contains the unit basis vectors and whose second and third rows contain the components of \mathbf{a} and \mathbf{b} . Then the cross product is obtained by following the usual rules for computing a 3×3 determinant.

$$\begin{aligned} \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} &= \mathbf{i} \det \begin{bmatrix} a_2 & a_3 \\ b_2 & b_3 \end{bmatrix} - \mathbf{j} \det \begin{bmatrix} a_1 & a_3 \\ b_1 & b_3 \end{bmatrix} + \mathbf{k} \det \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \\ &= [a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1] \end{aligned}$$

The geometric meaning of the cross product is given the following three properties:

1. $\mathbf{a} \times \mathbf{b}$ is orthogonal to \mathbf{a} and to \mathbf{b}
2. $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\|\|\mathbf{b}\|\sin(\theta)$, where θ is the angle between \mathbf{a} and \mathbf{b} . In this formula, θ lies between 0 and π , so that $\sin(\theta)$ is positive. This is the same as saying that the length of $\mathbf{a} \times \mathbf{b}$ is the area of the parallelogram spanned by \mathbf{a} and \mathbf{b} .
3. The vectors \mathbf{a} , \mathbf{b} and $\mathbf{a} \times \mathbf{b}$ obey the right hand rule.

This geometric description of the cross product shows that the definition of the cross product is independent of how we choose our co-ordinate axes. To verify 1, we compute the dot products $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b})$ and $\mathbf{b} \cdot (\mathbf{a} \times \mathbf{b})$ and verify that they are zero. (This is one of the problems below.)

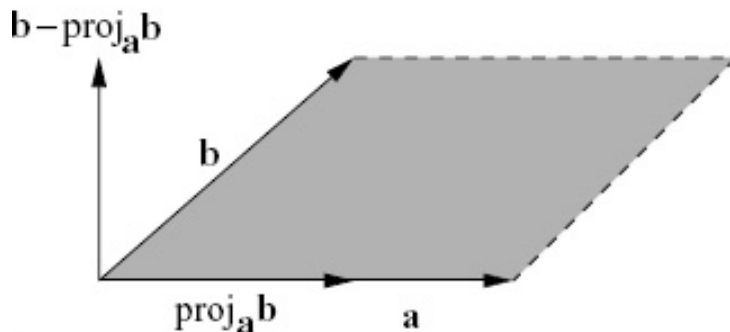
To verify 2 we must show that the length of $\mathbf{a} \times \mathbf{b}$ is the area of the parallelogram spanned by \mathbf{a} and \mathbf{b} , since the quantity $\|\mathbf{a}\|\|\mathbf{b}\|\sin(\theta)$ is precisely this area.

Since both the length and the area are positive quantities, it is enough to compare their squares. We have

$$\|\mathbf{a} \times \mathbf{b}\|^2 = (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2 \quad (2.4)$$

On the other hand, the area A of the parallelogram spanned by \mathbf{a} and \mathbf{b} is length $\|\mathbf{a}\|$ times the height. This height is the length of the vector $\mathbf{b} - \text{proj}_{\mathbf{a}} \mathbf{b} = \mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{a}/\|\mathbf{a}\|^2$ as shown in Figure 2.14. Using these facts, we arrive at the following formula for the square of the area of the parallelogram.

$$\begin{aligned} A^2 &= \|\mathbf{a}\|^2 \|\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{a}/\|\mathbf{a}\|^2\|^2 \\ &= \|\mathbf{a}\|^2 (\|\mathbf{b}\|^2 + (\mathbf{a} \cdot \mathbf{b})^2/\|\mathbf{a}\|^2 - 2(\mathbf{a} \cdot \mathbf{b})^2/\|\mathbf{a}\|^2) \\ &= \|\mathbf{a}\|^2 \|\mathbf{b}\|^2 - (\mathbf{a} \cdot \mathbf{b})^2 \\ &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2 \quad (2.5) \end{aligned}$$

Figure 2.14: The parallelogram spanned by \mathbf{a} and \mathbf{b} .

Expanding the expressions in (2.4) and (2.5) reveals that they are equal.

Notice that there are exactly two vectors satisfying properties 1 and 2, that is, perpendicular to the plane spanned by \mathbf{a} and \mathbf{b} and of a given length. The cross product of \mathbf{a} and \mathbf{b} is the one that satisfies the right hand rule. We say that vectors \mathbf{a} , \mathbf{b} and \mathbf{c} (the order is important) satisfy the right hand rule if you can point the index finger of your right hand in the direction of \mathbf{a} and the middle finger in the direction of \mathbf{b} and the thumb in the direction of \mathbf{c} . Try to convince yourself that if \mathbf{a} , \mathbf{b} and \mathbf{c} (in that order) satisfy the right hand rule, then so do \mathbf{b} , \mathbf{c} , \mathbf{a} and \mathbf{c} , \mathbf{a} , \mathbf{b} .

Here are some properties of the cross product that are useful in doing computations. The first two are maybe not what you expect.

1. $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
2. $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{c} \cdot \mathbf{a})\mathbf{b} - (\mathbf{b} \cdot \mathbf{a})\mathbf{c}$.
3. $s(\mathbf{a} \times \mathbf{b}) = (s\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (s\mathbf{b})$.
4. $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$.
5. $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$.

Example 2.11 Let $\mathbf{a} = (1, 3, -2)$ and $\mathbf{b} = (-1, 2, 3)$. Compute the following:

- (a) The area of the parallelogram whose sides are \mathbf{a} and \mathbf{b} .
- (b) The angle between \mathbf{a} and \mathbf{b} .

Solution:

- (a) The area of the parallelogram is equal to the length of $\mathbf{a} \times \mathbf{b}$:

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 3 & -2 \\ -1 & 2 & 3 \end{vmatrix} = \hat{i}(9 + 4) + \hat{j}(2 - 3) + \hat{k}(2 + 3) = (13, -1, 5)$$

so the area is

$$\|\mathbf{a} \times \mathbf{b}\| = \sqrt{13^2 + (-1)^2 + 5^2} = \sqrt{195} \approx 13.96$$

- (b) Note that the formula $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\|\|\mathbf{b}\|\sin\theta$ cannot be used for this question since it cannot distinguish between θ and $\pi - \theta$ (think about this point). Instead, use the cos, dot product formula which should always be used for the calculation of angles between vectors unless you really know what you are doing:

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|} = \frac{(1, 3, -2) \cdot (-1, 2, 3)}{\|(1, 3, -2)\|\|(-1, 2, 3)\|} = \frac{-1 + 6 - 6}{\sqrt{1+9+4}\sqrt{1+4+9}} = \frac{-1}{14}$$

so

$$\theta = \cos^{-1}\left(\frac{-1}{14}\right) \approx 1.64 \text{ radians or } \approx 94.10^\circ$$

Example 2.12 Consider the triangle T with three corners $(1, 1, 1)$, $(1, 2, 3)$ and $(2, 0, 1)$. Find the area of T . The area will be half of the area of the parallelogram spanned by (any) two distinct sides. We take sides $(1, 2, 3) - (1, 1, 1) = (0, 1, 2)$ and $(2, 0, 1) - (1, 1, 1) = (1, -1, 0)$ which make the computations a bit easier. We compute

$$(0, 1, 2) \times (1, -1, 0) = \det \begin{bmatrix} \hat{i} & \hat{j} & \hat{k} \\ 0 & 1 & 2 \\ 1 & -1 & 0 \end{bmatrix} = (2, 2, -1)$$

and then the area of the triangle is

$$1/2\|(2, 2, -1)\| = 3/2.$$

Try to convince yourself why it does not matter which two sides are taken in the computation above.

2.4.3 The triple product and the determinant in three dimensions

The cross product is defined so that the dot product of \mathbf{a} with $\mathbf{b} \times \mathbf{c}$ is a determinant:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \det \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

This determinant is called the triple product of \mathbf{a} , \mathbf{b} and \mathbf{c} .

Using this fact we can show that the absolute value of the triple product is the volume of the parallelepiped spanned by \mathbf{a} , \mathbf{b} and \mathbf{c} .

A diagram of the parallelepiped is shown in Figure 2.15. The absolute value of the triple product is

$$|\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})| = \|\mathbf{a}\| \cos(\theta) \|\mathbf{b} \times \mathbf{c}\|.$$

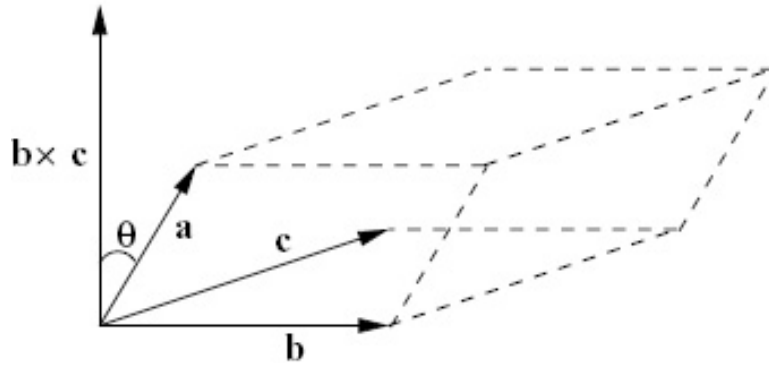


Figure 2.15: The Triple Product.

Here θ is the angle between \mathbf{a} and $\mathbf{b} \times \mathbf{c}$. The quantity $\|\mathbf{a}\| \cos(\theta)$ is the height of parallelepiped and $\|\mathbf{b} \times \mathbf{c}\|$ is the area of the base. The product of these is the volume of the parallelepiped, as claimed. Thus

$$\left| \det \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \right| = \text{Volume of the parallelepiped spanned by } \mathbf{a}, \mathbf{b} \text{ and } \mathbf{c}$$

The sign of the triple product is positive if θ is between zero and $\pi/2$ and negative if θ lies between $\pi/2$ and π . This is the case if \mathbf{a} is on the same side of the plane spanned by \mathbf{b} and \mathbf{c} as $\mathbf{b} \times \mathbf{c}$. This holds if the vectors \mathbf{b} , \mathbf{c} and \mathbf{a} (in that order) satisfy the right hand rule. Equivalently \mathbf{a} , \mathbf{b} and \mathbf{c} (in that order) satisfy the right hand rule.

Mathematically, it is more satisfactory to define the right hand rule using the determinant. That is, we say that vectors \mathbf{a} , \mathbf{b} and \mathbf{c} satisfy the right hand rule if the determinant $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is positive.

2.4.4 MATLAB: assigning matrices and det and cross commands

cross The command `cross(a,b)` computes the cross product $\mathbf{a} \times \mathbf{b}$. An error results if \mathbf{a} or \mathbf{b} are not vectors of length 3. As an example, the command

```
cross([1 0 0],[0 1 0])
```

gives the vector result `[0 0 1]`.

matrices: The syntax to generate a matrix is shown below using a 2×2 example

```
a = [1 2; 3 4]
```

This command assigns a matrix to `a` that has the vector `[1 2]` in its first row and `[3 4]` in its second. Entries of a matrix can be accessed individually, for example `a(1,2)` is the entry in the first row, second column.

zeros: Many applications can lead to large matrices with many rows and columns.

Even though MATLAB can do matrix computations, it can be tedious to enter these large matrices by hand. In some cases the matrices have mostly zeros as entries (these matrices are called *sparse*). In these cases it is more efficient to generate a matrix of all zeros and then modify the entries that are not zero. For example

```
a = zeros(2,2);  
a(1,1) = 1;
```

generates a 2×2 matrix with entries that are all zero except the upper left entry which is 1. Note that `zeros(n,m)` generates a matrix with n rows and m columns with all zero entries. So `zeros(1,m)` is a row vector of length m and `zeros(n,1)` is a column vector of length m with all zero entries.

rand: `rand (n,m)` generates a matrix with n rows and m columns with entries that are random numbers uniformly distributed in the interval $[0,1]$.

det: The command `det(a)` returns the determinant of the matrix `a`. An error occurs if `a` is not a square (same number of rows and columns) matrix. Determinants of larger matrices (than 2×2 and 3×3 discussed in this section) are discussed in Chapter 4.

2.4.5 MATLAB: generating scripts with the MATLAB editor

Often times using the command window in MATLAB to solve a problem can be tedious, because if the need arises to redo the problem, or change a parameter, one has to rewrite it all. The editor comes in handy for such cases. The editor is a text window (accessed from the command window: **File** \rightarrow **New** \rightarrow **Blank M-file**) where one can write commands in the same syntax as the editor, and when one runs it, the results appear in the command window exactly as if one had written them there one after the other.

For example, the code to generate three random orthogonal vectors would look something like this:

```
a1 = rand(3,1)  
b = rand(3,1);  
a2 = cross(a1,b)  
a3 = cross(a1,a2)  
dot(a1,a2)  
dot(a1,a3)  
dot(a2,a3)
```

Note that the last three lines are there to check that the three vectors are mutually orthogonal. Once the code was written, save it from the editor window: **File** → **Save as**, making sure that the name of the file has a “.m” extension (and the file name should contain no spaces). There are several different ways of running the script, the fastest one is to hit the F5 key. Alternatively, from the editor window it can be run from **Debug** → **Run**, or directly from the command window by typing the name of the script into the MATLAB command line.

2.4.6 MATLAB: floating point representation of real numbers

MATLAB can represent integers exactly (up to limited but large size). Using a “floating point representation”, MATLAB can represent most real numbers only approximately (but quite accurately - to 16 digits or so). In certain cases, the errors made in floating point approximation of numbers can be amplified and lead to noticeable errors in computed results. This will not happen typically in the examples and computer labs for Math 152, but the reader should be aware of the possibility.

Example 2.13 *Consider the vectors*

$$\begin{aligned}\mathbf{a} &= [1 \ 1 \ 1] \\ \mathbf{b} &= [\sqrt{2} \ \sqrt{2} \ 0]\end{aligned}$$

and $\mathbf{c} = \mathbf{a} + \mathbf{b}$. If a 3×3 matrix A is made with rows \mathbf{a} , \mathbf{b} and \mathbf{c} then the determinant of A is zero (by construction, the vectors lie on the same plane). If this computation is done in MATLAB,

```
a = [1 1 1];
b = [sqrt(2) sqrt(2) 0];
c = a + b;
A = [a; b; c];
det(A)
```

the result is $3.1402e-16$ not zero due to floating point approximation of intermediate computations. You can type `eps` in MATLAB to see the maximum relative error made by floating point approximation. On the computer used to do the computation above, `eps` had a value of $2.2204e-16$, so it is believable that the calculation error in the determinant was made by the combination of a few floating point approximations.

Try to determine how many decimal digits of accuracy the floating point representation in your calculator uses (typically, the accuracy is greater than what is displayed).

2.4.7 Problems

Problem 2.19 Compute $[1, 2, 3] \times [4, 5, 6]$

Problem 2.20 Use the definition to find the determinant of the matrix

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

Do the computation by hand showing your work, but you can check your result using MATLAB. From your result, decide if the vectors $[1 \ 1 \ 1]$, $[1 \ 2 \ 3]$ and $[1 \ 0 \ -1]$ lie in the same plane (justify your answer, very briefly).

Problem 2.21 Verify that $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) = 0$ and $\mathbf{b} \cdot (\mathbf{a} \times \mathbf{b}) = 0$.

Problem 2.22 Simplify each of the following expressions:

(a) $((1, 4, -1) \cdot (2, 1, 3))((2, 1, 4) \times (1, 4, 9))$

(b) $(7, 1, 0) \cdot ((2, 0, -1) \times (1, 4, 3))$

(c) $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{b} \times \mathbf{a})$

Problem 2.23 Explain why $\|\mathbf{a}\|\|\mathbf{b}\|\sin(\theta)$ is the area of the parallelogram spanned by \mathbf{a} and \mathbf{b} . Here θ is the angle between the two vectors.

Problem 2.24 Find examples to show that in general $\mathbf{a} \times \mathbf{b} \neq \mathbf{b} \times \mathbf{a}$ and $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.

Problem 2.25 (Matlab) The Matlab command $\mathbf{a}=\text{rand}(1,\mathbf{n})$ generates an $n \times 1$ vector with random entries. Write a script that generates three random vectors and write what you obtain from $\mathbf{a} \times \mathbf{b} - \mathbf{b} \times \mathbf{a}$, and from $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) - (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$. Does that constitute a proof?

Problem 2.26 Show that $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$.

Problem 2.27 (Matlab) Write a script that generates three random vectors and checks that the result from problem 2.26 holds: $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$.

Problem 2.28 Derive an expression for $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d})$ that involves dot products but not cross products.

Problem 2.29 (a) Draw a sketch containing the vectors \mathbf{a} , \mathbf{b} and $\mathbf{a} \times (\mathbf{a} \times \mathbf{b})$. Assume that \mathbf{a} and \mathbf{b} lie in the plane of the paper and have an acute angle between them.

(b) Find a formula for $\mathbf{a} \times (\mathbf{a} \times \mathbf{b})$ which involves only $\|\mathbf{a}\|$, \mathbf{b} and $\text{proj}_{\mathbf{a}}\mathbf{b}$.
Hint: use a property of the dot product.

Problem 2.30 What is the analog of the cross product in two dimensions? How about four dimensions?

2.5 Lines and Planes

2.5.1 Describing linear sets

The following sections we will consider points, lines, planes and space in two and three dimensions. Each of these sets have two complementary (or dual) descriptions. One is called the parametric form and the other the equation form. Roughly speaking, the parametric form specifies the set using vectors that are parallel to the set, while the equation form uses vectors that are orthogonal to the set. Using the parametric description, it is easy to write down explicitly all the elements of the set, but difficult to check whether a given point lies in the set. Using the equation description its the other way around: if someone gives you a point, it is easy to check whether it lies in the set, but it is difficult to write down explicitly even a single member of the set.

One way of thinking about solving a system of linear equations is simply going from one description to the other. This will (hopefully) become clear later on.

These sections will always follow the same pattern. We will consider the parametric and equation descriptions, first in the special case when the set passes through the origin. Then we consider the general case. Recall that when we say “the point \mathbf{x} ” this means “the point at the head of the vector \mathbf{x} whose tail is at the origin.” In these sections we have used the notation $[x_1, x_2]$ instead of $[x, y]$ and $[x_1, x_2, x_3]$ instead of $[x, y, z]$ for typical points in two and three dimensions.

2.5.2 Lines in two dimensions: Parametric form

First we consider lines passing through the origin. Let $\mathbf{a} = [a_1, a_2]$ be vector in the direction of the line. Then all the points \mathbf{x} on the line are of the form

$$\mathbf{x} = s\mathbf{a}$$

for some number s . The number s is called a parameter. Every value of s corresponds to exactly one point (namely $s\mathbf{a}$) on the line.

Now we consider the general case. Let \mathbf{q} be a point on the line and \mathbf{a} lie in the direction of the line. Then the points on the line can be thought of as the points on the line through the origin in the direction of \mathbf{a} shifted or translated by \mathbf{q} . Then a point \mathbf{x} lies on the line through \mathbf{q} in the direction of \mathbf{a} exactly when

$$\mathbf{x} = \mathbf{q} + s\mathbf{a}$$

for some value of s .

Example 2.14 Find a parametric form of the line that goes through the points $(1,2)$ and $(2,4)$. We find the direction vector from

$$(2, 4) - (1, 2) = (1, 2).$$

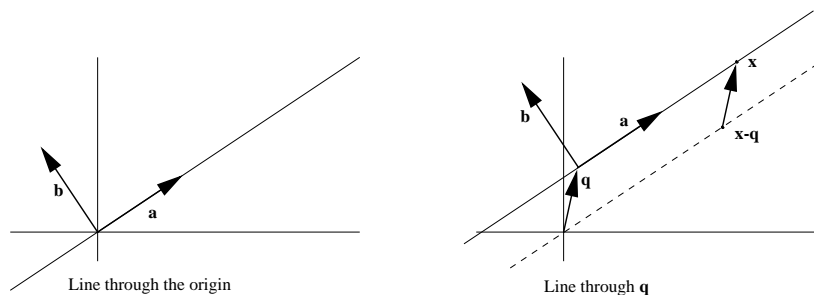


Figure 2.16: A line in two dimensions.

Now

$$(1, 2) + s(1, 2)$$

is a parametric form of the line. That is, every point on the line can be written in the form above with a (unique) value of s and every point of the form above is on the line. Note that parametric forms are not unique. The same line can be described by

$$(2, 4) + t(2, 4)$$

since we know the point $(2, 4)$ is on the line and the direction $(2, 4)$ is a scalar multiple of $(1, 2)$. Of course, specific points on the line will correspond to different values of s and t .

2.5.3 Lines in two dimensions: Equation form

First we consider lines passing through the origin shown in Figure 2.16 (left). Let $\mathbf{b} = [b_1, b_2]$ be orthogonal to the direction of line. The point \mathbf{x} is on the line exactly when $\mathbf{x} \cdot \mathbf{b} = 0$. This can be written

$$x_1 b_1 + x_2 b_2 = 0.$$

Now we consider the general case shown in Figure 2.16 (right). Let \mathbf{q} be a point on the line and $\mathbf{b} = [b_1, b_2]$ be orthogonal to the direction of line. A point \mathbf{x} lies on the line through \mathbf{q} in the direction of \mathbf{a} exactly when $\mathbf{x} - \mathbf{q}$ lies on the line through the origin in the direction of \mathbf{a} . Thus $(\mathbf{x} - \mathbf{q}) \cdot \mathbf{b} = 0$. This can be written

$$(x_1 - q_1)b_1 + (x_2 - q_2)b_2 = 0$$

or

$$x_1 b_1 + x_2 b_2 = c,$$

where $c = \mathbf{q} \cdot \mathbf{b}$.

Example 2.15 Find an equation form for the line in parametric form below:

$$(1, 2) + s(1, 2).$$

We need to find a vector perpendicular to $(1,2)$. We use the idea from section 2.4.1 and use

$$(1,2)^\perp = (-2,1)$$

Together with the point $(1,2)$ on the line, we have an equation form

$$-x_1 + 2x_2 = (-2,1) \cdot (1,2) = 0$$

Example 2.16 Find a parametric form for the line in equation form below:

$$x_1 + 4x_2 = 1. \tag{2.6}$$

We need to find a point (any point) on the line. It is simple to look for the x_2 intercept ($x_1 = 0$) and find the point $(1,0)$. The direction of the line will be perpendicular to $(1,4)$ so we will use

$$(1,4)^\perp = (-4,1)$$

to get the parametric form

$$(x_1, x_2) = (1,0) + s(-4,1)$$

Note that for every s we get a point that satisfies (2.6).

2.5.4 Lines in three dimensions: Parametric form

The parametric form of a line in three (or higher) dimensions looks just the same as in two dimensions. The points \mathbf{x} on the line are obtained by starting at some point \mathbf{q} on the line and then adding all multiples of a vector \mathbf{a} pointing in the direction of the line. So

$$\mathbf{x} = \mathbf{q} + s\mathbf{a}$$

The only difference is that now \mathbf{q} and \mathbf{a} are vectors in three dimensions.

Example 2.17 Find a parametric form for the line that passes through $(1,1,1)$ and $(1,2,3)$. Determine if the point $(1,-2,3)$ is on the line. We find the direction of the line

$$(1,2,3) - (1,1,1) = (0,1,2)$$

so a parametric form of the line is

$$(1,1,1) + s(0,1,2).$$

If $(1,-2,3)$ were on the line then

$$(1,1,1) + s(0,1,2) = (1,-2,3)$$

for some s . The above is a vector equation that must be satisfied for every component. The first component reads $1 + 0 \times s = 1$ which is true for all s . The second component reads $1 + s = -2$ which requires $s = -3$. If we now check the third component $1 - 3(2) = -5 \neq 3$. We conclude that $(1,-2,3)$ is not on the line.

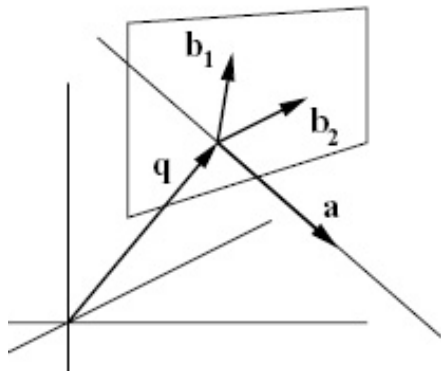


Figure 2.17: A line in three dimensions.

2.5.5 Lines in three dimensions: Equation form

We begin with lines through the origin. A line through the origin can be described as all vectors orthogonal to a plane. Choose two vectors \mathbf{b}_1 and \mathbf{b}_2 lying in the plane that are not collinear (or zero). Then a vector is orthogonal to the plane if and only if it is orthogonal to both \mathbf{b}_1 and \mathbf{b}_2 . Therefore the line consists of all points \mathbf{x} such that $\mathbf{x} \cdot \mathbf{b}_1 = 0$ and $\mathbf{x} \cdot \mathbf{b}_2 = 0$. If $\mathbf{b}_1 = [b_{1,1}, b_{1,2}, b_{1,3}]$ and $\mathbf{b}_2 = [b_{2,1}, b_{2,2}, b_{2,3}]$ then these equations can be written

$$\begin{aligned} b_{1,1}x_1 + b_{1,2}x_2 + b_{1,3}x_3 &= 0 \\ b_{2,1}x_1 + b_{2,2}x_2 + b_{2,3}x_3 &= 0 \end{aligned}$$

Notice that there are many possible choices for the vectors \mathbf{b}_1 and \mathbf{b}_2 . The method of Gaussian elimination, studied later in this course, is a method of replacing the vectors \mathbf{b}_1 and \mathbf{b}_2 with equivalent vectors in such a way that the equations become easier to solve.

Now consider a line passing through the point \mathbf{q} and orthogonal to the directions \mathbf{b}_1 and \mathbf{b}_2 as shown in Figure 2.17. A point \mathbf{x} lies on this line precisely when $\mathbf{x} - \mathbf{q}$ lies on the line through the origin that is orthogonal to \mathbf{b}_1 and \mathbf{b}_2 . Thus $(\mathbf{x} - \mathbf{q}) \cdot \mathbf{b}_1 = 0$ and $(\mathbf{x} - \mathbf{q}) \cdot \mathbf{b}_2 = 0$. This can be written

$$\begin{aligned} b_{1,1}(x_1 - q_1) + b_{1,2}(x_2 - q_2) + b_{1,3}(x_3 - q_3) &= 0 \\ b_{2,1}(x_1 - q_1) + b_{2,2}(x_2 - q_2) + b_{2,3}(x_3 - q_3) &= 0, \end{aligned}$$

or

$$\begin{aligned} b_{1,1}x_1 + b_{1,2}x_2 + b_{1,3}x_3 &= c_1 \\ b_{2,1}x_1 + b_{2,2}x_2 + b_{2,3}x_3 &= c_2 \end{aligned}$$

where $c_1 = \mathbf{q} \cdot \mathbf{b}_1$ and $c_2 = \mathbf{q} \cdot \mathbf{b}_2$

Example 2.18 Find a parametric and an equation form of the line that passes through the points $(0, 1, 5)$ and $(2, 2, 2)$. We find a direction vector

$$\mathbf{v} = (2, 2, 2) - (0, 1, 5) = (2, 1, -3)$$

and write a parametric form for the line

$$\mathbf{L} : \mathbf{x} = (0, 1, 5) + s(2, 1, -3).$$

To write an equation form we need to find two *different* (not collinear) vectors perpendicular to \mathbf{v} . To get one, we can use a variant of the procedure in Section 2.4.1 where we rotated 2D vectors by 90° to get a perpendicular vector.

$$\mathbf{b}_1 = (0, 3, 1).$$

You can check directly that \mathbf{b}_1 is perpendicular to \mathbf{v} ($\mathbf{b}_1 \cdot \mathbf{v} = 0$). To get the second vector \mathbf{b}_2 we can use the same idea on different components:

$$\mathbf{b}_2 = (-1, 2, 0).$$

Note that \mathbf{b}_2 is perpendicular to \mathbf{v} and not a multiple of \mathbf{b}_1 as required. An alternative for the second vector would be $\mathbf{b}_1 \times \mathbf{v}$ since that vector is perpendicular to both \mathbf{v} and \mathbf{b}_1 . With the formulas above, we can write an equation form for the line:

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

You can check your answer by showing that both of the original points satisfy both equations.

2.5.6 Planes in three dimensions: Parametric form

We begin with planes through the origin. Since a plane is a two dimensional object, we will need two parameters to describe points on the plane. Let \mathbf{a}_1 and \mathbf{a}_2 be non-collinear vectors in the direction of the plane. Then every point on the plane can be reached by adding some multiple of \mathbf{a}_1 to some other multiple of \mathbf{a}_2 . In other words, points \mathbf{x} on the plane are all points of the form

$$\mathbf{x} = s\mathbf{a}_1 + t\mathbf{a}_2$$

for some values of s and t .

If the plane passes through some point \mathbf{q} in the directions of \mathbf{a}_1 and \mathbf{a}_2 , then we simply shift all the points on the parallel plane through the origin by \mathbf{q} . So \mathbf{x} lies on the plane if

$$\mathbf{x} = \mathbf{q} + s\mathbf{a}_1 + t\mathbf{a}_2$$

for some values of s and t .

Example 2.19 Find a parametric form of the plane that passes through the points $(1, 0, 0)$, $(1, 1, 1)$ and $(1, 0, 2)$. We find direction vectors

$$\begin{aligned} \mathbf{a}_1 &= (1, 1, 1) - (1, 0, 0) = (0, 1, 1) \\ \mathbf{a}_2 &= (1, 0, 2) - (1, 0, 0) = (0, 0, 2). \end{aligned}$$

Note that \mathbf{a}_1 and \mathbf{a}_2 do not have the same direction (are not collinear). If they had been, then the three points would all be on the same line and so would not define a unique plane. We can proceed with the parametric form of the plane

$$\mathbf{x} = (1, 0, 0) + s\mathbf{a}_1 + t\mathbf{a}_2 = (1, s, s + 2t).$$

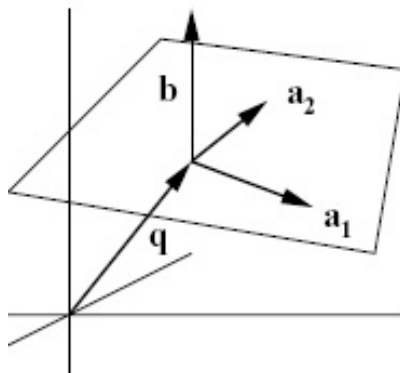


Figure 2.18: A plane in three dimensions.

2.5.7 Planes in three dimensions: Equation form

A plane through the origin can be described as all vectors orthogonal to a given vector \mathbf{b} as shown in Figure 2.18. (In this situation, if \mathbf{b} has unit length it is called the normal vector to the plane and is often denoted \mathbf{n} .) Therefore \mathbf{x} lies on the plane whenever $\mathbf{x} \cdot \mathbf{b} = 0$, or

$$b_1x_1 + b_2x_2 + b_3x_3 = 0.$$

If a plane with normal vector \mathbf{b} is translated so that it passes through the point \mathbf{q} , then \mathbf{x} lies on the plane whenever $\mathbf{x} - \mathbf{q}$ lies on the parallel plane through the origin. Thus \mathbf{x} lies on the plane whenever $(\mathbf{x} - \mathbf{q}) \cdot \mathbf{b} = 0$. Equivalently

$$b_1(x_1 - q_1) + b_2(x_2 - q_2) + b_3(x_3 - q_3) = 0,$$

or

$$b_1x_1 + b_2x_2 + b_3x_3 = c,$$

where $c = \mathbf{q} \cdot \mathbf{b}$.

Example 2.20 Write the equation form of the plane described by the parametric form below:

$$\mathbf{x} = (1 + s + t, 5t, -2)$$

Note: It is possible to see the answer from the form above, but let us go through the steps of the calculations you would do if the example were more complicated. Write the expression above as

$$\mathbf{x} = (1, 0, -2) + s(1, 0, 0) + t(1, 5, 0).$$

In this form we can recognize $\mathbf{q} = (1, 0, -2)$ as a point on the plane (corresponding to parameters $s = 0$ and $t = 0$) and the vectors $\mathbf{a}_1 = (1, 0, 0)$ and $\mathbf{a}_2 = (1, 5, 0)$ as directions *in* the plane. To write the equation form of the

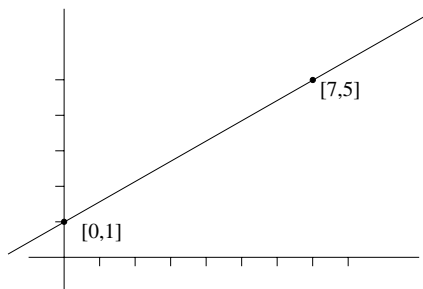


Figure 2.19: Diagram for problem 2.32.

plane, we need a point on it (which we have) and the normal direction \mathbf{b} to it, which must be perpendicular to \mathbf{a}_1 and \mathbf{a}_2 . Thus we can compute

$$\mathbf{b} = \mathbf{a}_1 \times \mathbf{a}_2 = (0, 0, 5).$$

Now the equation form of the plane is $\mathbf{b} \cdot \mathbf{x} = \mathbf{b} \cdot \mathbf{q}$:

$$0x + 0y + 5z = 0 \cdot 1 + 0 \cdot 1 + 5 \cdot -2 = -10 \quad \text{or } z = -2$$

where in this case we have written the components of the vector \mathbf{x} as (x, y, z) .

Example 2.21 Find a parametric form for the plane $x + y + 2z = 2$. We identify the normal direction to the plane $\mathbf{b} = (1, 1, 2)$ by inspection. We need a point on the plane. It is easy to find one on the coordinate axes, $(2, 0, 0)$ for example. To get two directions \mathbf{a}_1 and \mathbf{a}_2 in the plane, we need two different directions perpendicular to \mathbf{b} . Proceeding as in Example 2.18 we take

$$\mathbf{a}_1 = (0, -2, 1) \quad \text{and} \quad \mathbf{a}_2 = (-1, 1, 0)$$

leading to a parametric form

$$\mathbf{x} = (1, 0, 0) + s\mathbf{a}_1 + t\mathbf{a}_2 = (1 - t, -2s + t, s).$$

2.5.8 Problems

Problem 2.31 A line orthogonal to \mathbf{b} can be described as the set of all points \mathbf{x} whose projections onto \mathbf{b} all have the same value. Using the formula for projections, show that this leads to the equation description of the line.

Problem 2.32 Find both the parametric form and equation form for the line in Figure 2.19. Write down five points on the line (notice that the parametric form is more useful for this). Check whether the point $[\frac{1012}{3}, \frac{1069}{21}]$ is on the line (notice that the equation form is more useful for this.)

Problem 2.33 Find the equation form for the line $[1, 1] + s[-1, 2]$.

Problem 2.34 Find the parametric form for the line $x_1 - 3x_2 = 5$

Problem 2.35 Use a projection to find the distance from the point $[-2, 3]$ to the line $3x_1 - 4x_2 = -4$

Problem 2.36 Consider the plane $x - y + 2z = 7$.

1. What is the normal direction to the plane?
2. Find the coordinates of any point (your choice) on the plane.

Problem 2.37 Let \mathbf{a} , \mathbf{b} and \mathbf{c} be the vertices of a triangle. By definition, the median of a triangle is a straight line that passes through a vertex of the triangle and through the midpoint of the opposite side.

- (i) Find the parametric form of the equation for each median.
- (ii) Do all the medians meet at a common point? If so, which point?

Problem 2.38 Find a pair of equations which define the line

$$\{(2, 0, -4) + s(0, 1, 3) : s \in \mathbb{R}\}$$

Problem 2.39 Find the intersection point between the line

$$\{(2, -1, 6) + s(1, -1, 0) : s \in \mathbb{R}\}$$

and the plane

$$\{t(0, 1, -3) + u(-1, 2, 0) : t, u \in \mathbb{R}\}$$

Hint: first find an equation for the plane.

Problem 2.40 Find the intersection point of the line with parametric form below

$$(1, 2, 3) + t(1, 0, -1)$$

and the plane

$$x + 2y - z = 5.$$

Problem 2.41 Find the equation of the plane containing the points $[1, 0, 1]$, $[1, 1, 0]$ and $[0, 1, 1]$.

Problem 2.42 Find the equation of the sphere which has the two planes $x_1 + x_2 + x_3 = 3$ and $x_1 + x_2 + x_3 = 9$ as tangent planes if the centre of the sphere is on the planes $2x_1 - x_2 = 0$, $3x_1 - x_3 = 0$.

Problem 2.43 The planes $x + y + z = 2$ and $x - y + 2z = 7$ intersect in a line. Find a parametric representation of this line.

Problem 2.44 Find the equation of the plane that passes through the point $[-2, 0, 1]$ and through the line of intersection of $2x_1 + 3x_2 - x_3 = 0$, $x_2 - 4x_2 + 2x_3 = -5$.

Problem 2.45 What's wrong with the question "Find the equation for the plane containing $[1, 2, 3]$, $[2, 3, 4]$ and $[3, 4, 5]$."?

Problem 2.46 Find the distance from the point \mathbf{p} to the plane $\mathbf{b} \cdot \mathbf{x} = c$.

Problem 2.47 Find the equation for the line through $[2, -1, -1]$ and parallel to each of the two planes $x_1 + x_2 = 0$ and $x_1 - x_2 + 2x_3 = 0$. Express the equation for the line in both parametric and equation form.

Problem 2.48 (Matlab) Plotting figures in Matlab is quite simple. Simply type `plot(1,2)` in the command window and observe what you get. Copy the following script (call it `basicplot.m` for example), and run it:

```
x = -2:0.1:2
m = 2
x_0 = 1
y_0 = 1
y = y_0 + m*(x-x_0)
plot(x,y, 'r')
```

Type `help plot` in the command window to learn about different options for the `plot` command. Without closing the figure window, type `hold on` in the command window, and re-run the script after changing the slope to `m = -1/2` (the `hold on` command allows you to overlap plots). Notice that the two lines should be perpendicular, but because of the scaling they appear not to be. Type in the command `axis equal` to fix that. How would you modify the script to plot a circle of radius 3?

2.6 Introduction to Linear Systems

2.6.1 Description of points and the geometry of solutions to systems of equations

So far we have considered the parametric and equation descriptions of lines and planes in two and three dimensions. We can also try to describe points in the same way. This will help you get a geometric picture of what it means to solve a system of equations.

The "parametric" description of a point doesn't have any parameters! It simply is the name of the point $\mathbf{x} = \mathbf{q}$. In two dimensions the equation form for describing a point will look like

$$\begin{aligned} b_{1,1}x_1 + b_{1,2}x_2 &= c_1 \\ b_{2,1}x_1 + b_{2,2}x_2 &= c_2 \end{aligned}$$

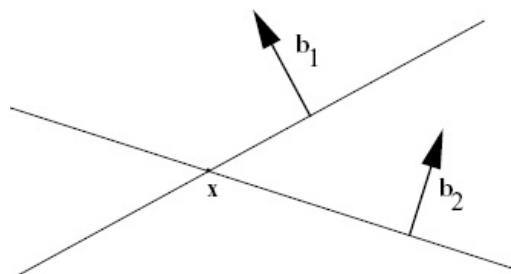


Figure 2.20: Intersection of lines in 2D that are not collinear is a point.

where the vectors $\mathbf{b}_1 = [b_{1,1}, b_{1,2}]$ and $\mathbf{b}_2 = [b_{2,1}, b_{2,2}]$ are not collinear (not multiples of each other). Each equation describes a line. The point $\mathbf{x} = [x_1, x_2]$ will satisfy both equations if it lies on both lines, i.e., on the intersection. Since the vectors \mathbf{b}_1 and \mathbf{b}_2 are not co-linear, the lines are not parallel, so the intersection is a single point. This situation is shown in Figure 2.20.

In three dimensions the equation form for describing a point will look like

$$\begin{aligned} b_{1,1}x_1 + b_{1,2}x_2 + b_{1,3}x_3 &= c_1 \\ b_{2,1}x_1 + b_{2,2}x_2 + b_{2,3}x_3 &= c_2 \\ b_{3,1}x_1 + b_{3,2}x_2 + b_{3,3}x_3 &= c_3 \end{aligned}$$

where \mathbf{b}_1 , \mathbf{b}_2 and \mathbf{b}_3 don't all lie on the same plane. This can be interpreted as the intersection of three planes in a single point.

Notice that going from the equation description of a point to the parametric description just means finding the solution of the system of equations. If, in two dimensions, the vectors \mathbf{b}_1 and \mathbf{b}_2 are not collinear, or in three dimensions, \mathbf{b}_1 , \mathbf{b}_2 and \mathbf{b}_3 don't all lie on the same plane, then the system of equations has a unique solution.

Now suppose that you are handed an arbitrary system of equations

$$\begin{aligned} b_{1,1}x_1 + b_{1,2}x_2 + b_{1,3}x_3 &= c_1 \\ b_{2,1}x_1 + b_{2,2}x_2 + b_{2,3}x_3 &= c_2 \\ b_{3,1}x_1 + b_{3,2}x_2 + b_{3,3}x_3 &= c_3 \end{aligned}$$

What does the set of solutions $\mathbf{x} = [x_1, x_2, x_3]$ look like? As we just have seen, if \mathbf{b}_1 , \mathbf{b}_2 and \mathbf{b}_3 don't all lie on the same plane, there is a unique solution given as the intersection of three planes. Recall that the determinant can be used to test whether the vectors \mathbf{b}_1 , \mathbf{b}_2 and \mathbf{b}_3 lie on the same plane. So a unique solution exists to the equation precisely when

$$\det \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{bmatrix} \neq 0$$

What happens when the determinant is zero and three vectors \mathbf{b}_1 , \mathbf{b}_2 and \mathbf{b}_3 do lie on the same plane? Then it could be that the three planes intersect in a line.

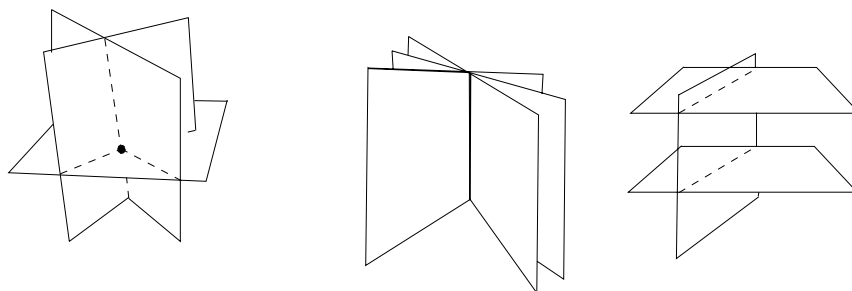


Figure 2.21: Planes intersecting.

In this case every point on that line is a solution of the system of equations, and the solution set has a parametric description of the form $\mathbf{x} = \mathbf{q} + s\mathbf{a}$. It could also be that all three planes are the same, in which case the solution set is the plane. In this case the solution set has a parametric description of the form $\mathbf{x} = \mathbf{q} + s_1\mathbf{a}_1 + s_2\mathbf{a}_2$. Another possibility is that two of the planes could be parallel with no intersection. In this case there are no solutions at all! Some of these possibilities are illustrated in Figure 2.21.

Example 2.22 Determine whether the following system of linear equations has a solution. It is not necessary to find the solution or solutions if there are any.

$$\begin{aligned}x_1 + x_2 + x_3 &= 3 \\x_1 + \quad + 2x_3 &= 3 \\x_1 + 2x_2 + x_3 &= 4\end{aligned}$$

We compute the determinant

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

Since the determinant is not zero, we know there is a unique solution point. By inspection, we can see that $(1,1,1)$ solves all three equations and since we know this solution is unique, it is the only point that satisfies all three equations. In Chapter 3 we will learn a systematic way to find the solution or all solutions (if any) of any linear systems.

Example 2.23 Determine whether the following system of linear equations has a solution. It is not necessary to find the solution or solutions if there are any.

$$\begin{aligned}x_1 + x_2 + x_3 &= 1 \\x_1 + \quad + x_3 &= 2 \\x_1 + 2x_2 + x_3 &= 1\end{aligned}$$

We compute the determinant

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

Now we know that the system does not have a unique solution. Consider Figure 2.21. It can be made rigorous that there are two cases that could come from this system: either the solution describes a line (middle picture, infinite solutions) or the three planes do not intersect (right picture, no solutions). In Chapter 3 we will develop an algorithm to decide this question. Here, we can look at combinations of the equations. Take twice the first equation and subtract the second to get

$$x_1 + 2x_2 + x_3 = 0.$$

Since this contradicts the third equation, we see that this system has no solutions.

2.6.2 Describing the whole plane in two dimensions and all of space in three dimensions

If the set we are trying to describe is the whole plane in two dimensions or all of space in three dimensions, then we don't need any equations, since there are no restrictions on the points. However it does make sense to think about the parametric form.

Lets start with two dimensions. Consider Figure 2.22. If we pick any two vectors \mathbf{a}_1 and \mathbf{a}_2 that don't lie on the same line (that is they are different directions, not collinear), then any vector $\mathbf{x} = [x_1, x_2]$ in the plane can be written as $s_1\mathbf{a}_1 + s_2\mathbf{a}_2$. Notice that every choice of s_1 and s_2 corresponds to exactly one vector \mathbf{x} . In this situation we could use the parameters s_1 and s_2 as co-ordinates instead of x_1 and x_2 . In fact if \mathbf{a}_1 and \mathbf{a}_2 are unit vectors orthogonal to each other, this just amounts to changing the co-ordinate axes to lie along \mathbf{a}_1 and \mathbf{a}_2 . The new co-ordinates $[s_1, s_2]$ are then just what we were calling $[x'_1, x'_2]$ before. In fact, even if the vectors \mathbf{a}_1 and \mathbf{a}_2 are not unit vectors orthogonal to each other, we can still think of them of lying along new co-ordinate axes. However, now the axes have been stretched and sheared instead of just rotated, and need not lie at right angles any more. The discussion above motivates two important definitions.

Definition 2.1 *A combination of vectors of the form*

$$s_1\mathbf{a}_1 + s_2\mathbf{a}_2$$

is called a linear combination of $\{\mathbf{a}_1, \mathbf{a}_2\}$. The definition includes underlying sets of more vectors. That is,

$$s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + s_3\mathbf{a}_3$$

is a linear combination of $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ and so on.

Writing vectors as linear combinations of a set of vectors has the interpretation of expressing the vector in a new coordinate system.

Definition 2.2 *The set of all linear combinations of a set of vectors is called the span of the set of vectors.*

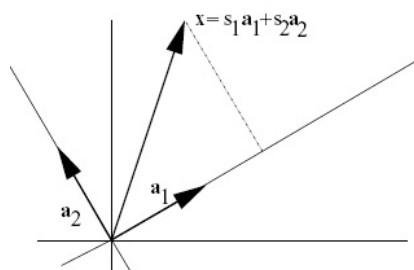


Figure 2.22: A basis in 2D

The situation in three dimensions is similar. Now we must pick three vectors \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 that don't lie on the same plane. Then every vector \mathbf{x} has a unique representation $\mathbf{x} = s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + s_3\mathbf{a}_3$. We can say that the span of $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ is all of \mathbb{R}^3 . Again, we could use s_1 , s_2 and s_3 as co-ordinates in place of x_1 , x_2 and x_3 . Again, if \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 are orthogonal with unit length, then this amounts to choosing new (orthogonal) co-ordinate axes.

2.6.3 Linear dependence and independence

The condition in two dimensions that two vectors are not co-linear, and the condition in three dimensions that three vectors do not lie on the same plane has now come up several times — in ensuring that a system of equations has a unique solutions and in ensuring that every vector can be written in a unique way as a linear combination of those vectors. This condition can be tested by computing a determinant.

We will now give this condition a name and define the analogous condition in any number of dimensions. Recall the definition that if $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ is a collection of vectors then a vector of the form

$$s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + \dots + s_n\mathbf{a}_n$$

for some choice of numbers s_1, \dots, s_n is called a *linear combination* of $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$.

Definition 2.3 A collection of vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ is called linearly dependent if some linear combination of them equals zero, i.e.,

$$s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + \dots + s_n\mathbf{a}_n = \mathbf{0}$$

for s_1, \dots, s_n not all zero. A collection of vectors is said to be linearly independent if it is not linearly dependent. In other words, the vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ are linearly independent if the only way a linear combination of them $s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + \dots + s_n\mathbf{a}_n$ can equal zero is for $s_1 = s_2 = \dots = s_n = 0$.

What does linear dependence mean in three dimensions? Suppose that \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 are linearly dependent. Then there are some numbers s_1 , s_2 and s_3 ,

not all zero, such that

$$s_1 \mathbf{a}_1 + s_2 \mathbf{a}_2 + s_3 \mathbf{a}_3 = \mathbf{0}.$$

Suppose that s_1 is one of the non-zero numbers. Then we can divide by $-s_1$ and find that

$$-\mathbf{a}_1 + s'_2 \mathbf{a}_2 + s'_3 \mathbf{a}_3 = \mathbf{0}$$

for $s'_2 = -s_2/s_1$ and $s'_3 = -s_3/s_1$. Thus

$$\mathbf{a}_1 = s'_2 \mathbf{a}_2 + s'_3 \mathbf{a}_3,$$

or \mathbf{a}_1 is a linear combination of \mathbf{a}_2 and \mathbf{a}_3 . But this implies that \mathbf{a}_1 lies on the plane spanned by \mathbf{a}_2 and \mathbf{a}_3 , i.e., the vectors all lie on the same plane. If s_1 happens to be zero we can repeat the same argument with one of the s_i 's which is not zero. Thus linear dependence implies that all three vectors lie on the same plane. Conversely, if all three vectors lie on the same plane, then we can write one vector as a linear combination of the other two, $\mathbf{a}_1 = s'_2 \mathbf{a}_2 + s'_3 \mathbf{a}_3$ which implies $-\mathbf{a}_1 + s'_2 \mathbf{a}_2 + s'_3 \mathbf{a}_3 = \mathbf{0}$ which says that the vectors are linearly dependent.

So in three dimensions, linear dependence means the vectors lie on the same plane. Similarly, in two dimensions, linear dependence means the vectors are co-linear.

Definition 2.4 A collection of n linearly independent vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ in \mathbb{R}^n dimensional space is called a basis. If $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ is a basis, then every vector \mathbf{x} can be written in a unique way as a linear combination

$$\mathbf{x} = s_1 \mathbf{a}_1 + s_2 \mathbf{a}_2 + \dots + s_n \mathbf{a}_n$$

Two linearly independent vectors in 2D is a basis for \mathbb{R}^2 . Three linearly independent vectors in 3D is a basis for \mathbb{R}^3 . The span of two linearly independent vectors in \mathbb{R}^3 is a plane through the origin. This motivates the following:

Definition 2.5 The dimension of a span of a set \mathcal{C} of linearly independent vectors is equal to the number of vectors in \mathcal{C} .

Example 2.24 Show that $(1, 1)$ and $(2, 0)$ are linearly independent. If they were linearly dependent by the definition they would be multiples of one another, which they are not. Thus, the vectors are linearly independent.

Example 2.25 Write $(3, 4)$ as a linear combination of $(1, 1)$ and $(2, 0)$. From the previous example, we know that $(1, 1)$ and $(2, 0)$ are linearly independent, and so are a basis for \mathbb{R}^2 and so any vector in \mathbb{R}^2 can be written uniquely as a linear combination of them. We write

$$(3, 4) = s(1, 1) + t(2, 0)$$

or

$$\begin{aligned} s + 2t &= 3 \\ s + 0t &= 4 \end{aligned}$$

It can be seen that writing a vector as a linear combination results in a linear system of equations. We will come back to this in Chapter 3. From the second component (second equation above), we see that $s = 4$. Going to the first equation we can then determine $t = -1/2$. Thus

$$(3, 4) = 4(1, 1) - \frac{1}{2}(2, 0)$$

Example 2.26 Show that $\{(1, 1, 1), (1, 1, 2), (1, 2, 1)\}$ is a linearly independent set of vectors. We compute

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

Thus, the set of vectors is linearly independent.

Example 2.27 What is the dimension of the span of $\mathcal{C} = \{(1, 2, 3, 0), (1, 1, 1, 1)\}$? By the definition of linear independence, two vectors can only be linearly dependent if they are multiples of each other. Thus, \mathcal{C} is a linearly independent set and its span is two dimensional. It is a two dimensional set that includes the origin in \mathbb{R}^4 .

2.6.4 Problems

Problem 2.49 Is the collection of vectors $\mathbf{a}_1 = [1, 1]$, $\mathbf{a}_2 = [1, 0]$ a basis for two dimensional space? If so, express the vector $\mathbf{x} = [0, 1]$ as a linear combination of \mathbf{a}_1 and \mathbf{a}_2

Problem 2.50 Let $\mathbf{a} = [2, 2, 2]$ and $\mathbf{b} = [3, 4, 1]$. Find all vectors \mathbf{c} such that the list of vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is not a basis of \mathbb{R}^3 .

Problem 2.51 Show that the collection of vectors $\mathbf{a} = [1, 1, 1]$, $\mathbf{b} = [1, 1, 0]$ and $\mathbf{c} = [1, 0, 0]$ is a basis of \mathbb{R}^3 . Express $[1, 2, 3]$ as a linear combination of the vectors \mathbf{a} , \mathbf{b} and \mathbf{c} .

Problem 2.52 Let the vectors $\mathbf{a} = [1, 0, 4]$, $\mathbf{b} = [2, -1, 0]$ and $\mathbf{c} = [8, -3, 8]$. Do these vectors form a basis of \mathbb{R}^3 ?

Problem 2.53 Is it possible for four vectors to be linearly independent in three dimensional space?

Problem 2.54 Show that the collection of vectors $\mathbf{a} = [2, 1, 3]$, $\mathbf{b} = [1, 0, 2]$ and $\mathbf{c} = [3, 0, 0]$ is a basis of \mathbb{R}^3 . Express $[12, 2, 4]$ as a linear combination of the vectors \mathbf{a} , \mathbf{b} and \mathbf{c} .

Problem 2.55 Suppose that $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ is a basis. Show that if some vector \mathbf{x} has representation $\mathbf{x} = s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + \dots + s_n\mathbf{a}_n$ and $\mathbf{x} = t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \dots + t_n\mathbf{a}_n$, then $s_1 = t_1, s_2 = t_2, \dots, s_n = t_n$. (Hint: subtract the two expressions for \mathbf{x} and use the fact that the basis vectors are linearly independent.)

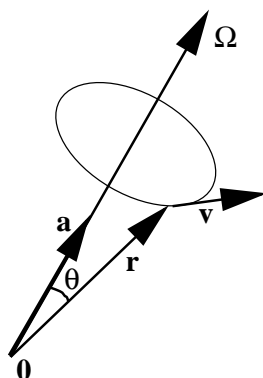


Figure 2.23: Rotational Motion.

Problem 2.56 (Matlab) Start by convincing yourself that the vectors $(1, 3)$, and $(5, 2)$ are linearly independent. This can easily be shown by writing a script that plots the vectors:

```
plot([1,0],[3,0])
hold on
plot([5,0],[2,0])
```

We can extend this script by also plotting a linear combination of the two vectors:

```
alfa=1;
beta=1;
clf()
plot([alfa*1,0],[alfa*3,0])
hold on
plot([beta*5,0],[beta*2,0])
plot([alfa*1+beta*5,alfa*1],[alfa*3+beta*2,alfa*3], '--')
```

Play around with the script and find parameters alfa and beta so that the vector $(1, 1)$ can be written as a linear combination of $(1, 3)$ and $(5, 2)$.

2.7 Additional Topics

These topics are not covered in the current version of Math 152 at UBC.

2.7.1 Application: rotational motion

Consider a rigid body rotating about an axis given by the unit vector \mathbf{a} at a rate of Ω radians per second. Let \mathbf{r} be the position vector of a point on the body as shown in Figure 2.23.

What is the velocity of the point? The point travels on a circle with radius $\|\mathbf{r}\| \sin(\theta)$, where θ is the angle that \mathbf{r} makes with the axis. Therefore, in one second, the point travels a distance of $\Omega\|\mathbf{r}\| \sin(\theta)$. Thus

- (i) the magnitude of the velocity is $\|\mathbf{v}\| = \Omega\|\mathbf{r}\| \sin(\theta)$.
- (ii) Now notice that \mathbf{v} is orthogonal to the plane spanned by \mathbf{a} and \mathbf{r} .
- (iii) Finally notice that $\Omega\mathbf{a}$, \mathbf{r} and \mathbf{v} obey the right hand rule.

The facts (i), (ii) and (iii) imply that \mathbf{v} is exactly the cross product of $\Omega\mathbf{a}$ and \mathbf{r} . It is customary to let $\boldsymbol{\Omega}$ denote the vector $\Omega\mathbf{a}$. Then

$$\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{r}.$$

Problem 2.57 *A body rotates at an angular velocity of 10 rad/sec about the axis through the points $[1, 1, -1]$ and $[2, -3, 1]$. Find the velocity of the point $[1, 2, 3]$ on the body.*

Problem 2.58 *The line L passing through the origin and the point $[1, 1, 1]$ passes through the midpoint of a thin metal rod of length 2 that is oriented in the direction $[1, 0, 0]$. The rod begins rotating about L at 3 revolutions per minute. What is the fastest speed (length of velocity vector) of any point on the rod?*

Problem 2.59 *Imagine a plate that lies in the xy -plane and is rotating about the z -axis. Let P be a point that is painted on this plane. Denote by r the distance from P to the origin, by $\theta(t)$ the angle at time t between the line from the origin to P and the x -axis and by $[x(t), y(t)]$ the co-ordinates of P at the time t . Find $x(t)$ and $y(t)$ in terms of $\theta(t)$. Compute the velocity of P in two ways: 1. by differentiating $[x(t), y(t)]$ and 2. by computing $\boldsymbol{\Omega} \times \mathbf{r}$.*

2.7.2 Application: 3-D graphics

How can we represent a three dimensional object on piece of paper or computer screen? Imagine the object in space and, a certain distance away, a point \mathbf{p} representing the eye of the observer. Between the observer and the object is a plane called the view plane. The position of the origin of this plane is described by a point \mathbf{q} , and its orientation is given by three orthogonal unit vectors of length 1, denoted \mathbf{e}_1 , \mathbf{e}_2 and \mathbf{e}_3 . (These are not the same as the standard basis vectors \mathbf{i} , \mathbf{j} and \mathbf{k} in this problem.) This situation is shown in Figure 2.24. As usual, only the heads of the vectors (points) \mathbf{p} , \mathbf{x} , \mathbf{y} and \mathbf{q} are shown on the diagram. (The origin, where the tails of these vectors lie, is not depicted at all.) We will assume that the view plane is a distance one from the observer in the direction \mathbf{e}_3 . Thus, \mathbf{e}_3 can be thought of as the direction that the observer is looking. Think of light rays leaving the object at point \mathbf{x} and travelling to the observer's eye at \mathbf{p} . At some point \mathbf{y} this line intersects the view plane. All the vectors \mathbf{y} on the view plane that correspond to some vector \mathbf{x} on our object will furnish the two dimensional representation of the object.

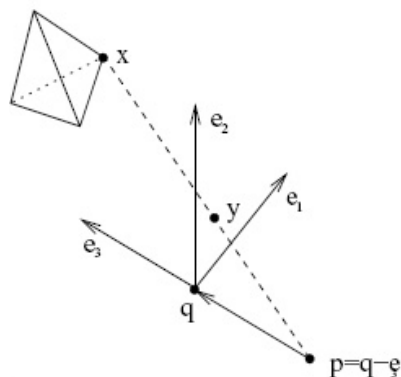


Figure 2.24: Perspective in 3-D graphics application.

How do we determine the point \mathbf{y} ? The parametric form of points on the plane is $\mathbf{q} + s_1\mathbf{e}_1 + s_2\mathbf{e}_2$. So we must have that $\mathbf{y} = \mathbf{q} + s_1\mathbf{e}_1 + s_2\mathbf{e}_2$ for some values of s_1 and s_2 . We also know that the vector $\mathbf{x} - \mathbf{p}$ is in the same direction as $\mathbf{y} - \mathbf{p}$. Therefore they must be multiples, i.e., $\mathbf{y} - \mathbf{p} = \lambda(\mathbf{x} - \mathbf{p})$ for some number λ . Substituting in our expression for \mathbf{y} yields

$$\mathbf{q} + s_1\mathbf{e}_1 + s_2\mathbf{e}_2 - \mathbf{p} = \lambda(\mathbf{x} - \mathbf{p}).$$

Since $\mathbf{q} - \mathbf{p} = \mathbf{e}_3$ this gives

$$\mathbf{e}_3 + s_1\mathbf{e}_1 + s_2\mathbf{e}_2 = \lambda(\mathbf{x} - \mathbf{p}).$$

Let us take the dot product of both sides of this equation with the unit vectors \mathbf{e}_3 , \mathbf{e}_1 and \mathbf{e}_2 . We can use the fact that $\mathbf{e}_i \cdot \mathbf{e}_j$ is zero if $i \neq j$ and 1 if $i = j$. Start with \mathbf{e}_3 . This gives

$$\mathbf{e}_3 \cdot \mathbf{e}_3 = 1 = \lambda \mathbf{e}_3 \cdot (\mathbf{x} - \mathbf{p}).$$

This determines λ .

$$\lambda = \frac{1}{\mathbf{e}_3 \cdot (\mathbf{x} - \mathbf{p})}.$$

Now take the dot product with \mathbf{e}_1 . This gives

$$s_1 = \lambda \mathbf{e}_1 \cdot (\mathbf{x} - \mathbf{p}) = \frac{\mathbf{e}_1 \cdot (\mathbf{x} - \mathbf{p})}{\mathbf{e}_3 \cdot (\mathbf{x} - \mathbf{p})}$$

Similarly, taking the dot product with \mathbf{e}_2 leads to

$$s_2 = \frac{\mathbf{e}_2 \cdot (\mathbf{x} - \mathbf{p})}{\mathbf{e}_3 \cdot (\mathbf{x} - \mathbf{p})}$$

To plot the image of an object, we now simply plot the co-ordinates s_1 and s_2 corresponding to all the points on the object on the s_1 - s_2 plane.

Example 2.28 Take $\mathbf{p} = [11, 0, 0]$, $\mathbf{q} = [10, 0, 0]$, $\mathbf{e}_1 = [0, 1, 0]$, $\mathbf{e}_2 = [0, 0, 1]$ and $\mathbf{e}_3 = [-1, 0, 0]$. What is the image of the point $\mathbf{x} = [1, 1, 1]$? We compute $\mathbf{x} - \mathbf{p} = [-10, 1, 1]$ so that

$$\begin{aligned}\mathbf{e}_1 \cdot (\mathbf{x} - \mathbf{p}) &= 1 \\ \mathbf{e}_2 \cdot (\mathbf{x} - \mathbf{p}) &= 1 \\ \mathbf{e}_3 \cdot (\mathbf{x} - \mathbf{p}) &= 10\end{aligned}$$

So $s_1 = s_2 = 1/10$.

Example 2.29 Continue the previous example and compute the image of a line segment between $[1, 1, 1]$ and $[2, 0, 1]$. These are all points of the form $\mathbf{x} = [1, 1, 1] + t([2, 0, 1] - [1, 1, 1]) = [1 + t, 1 - t, 1]$ as t varies between 0 and 1. This time we have $\mathbf{x} - \mathbf{p} = [1 + t - 11, 1 - t, 1]$ so that

$$\begin{aligned}\mathbf{e}_1 \cdot (\mathbf{x} - \mathbf{p}) &= 1 - t \\ \mathbf{e}_2 \cdot (\mathbf{x} - \mathbf{p}) &= 1 \\ \mathbf{e}_3 \cdot (\mathbf{x} - \mathbf{p}) &= 10 - t\end{aligned}$$

Thus $s_1 = (1-t)/(10-t)$ and $s_2 = 1/(10-t)$. Even though it is not immediately obvious, the points $[s_1, s_2]$, as t varies, all lie on a line segment. In fact

$$s_1 + 9s_2 = (1-t)/(10-t) + 9/(10-t) = (10-t)/(10-t) = 1.$$

This shows that the points $[s_1, s_2]$ lie on a line perpendicular to $[1, 9]$.

In fact, it is possible to show that any line segment in space maps to a line segment on the s_1 - s_2 plane. Thus, to plot the image of an object consisting of straight line segments (such as the tetrahedron in the picture) it is only necessary to plot the vertices and then join them by straight lines.

Problem 2.60 What are the s_1 and s_2 co-ordinates of the point $\mathbf{x} = [1, 2, 3]$, if \mathbf{p} , \mathbf{q} are as above, $\mathbf{e}_1 = [0, 1/\sqrt{2}, 1/\sqrt{2}]$ and $\mathbf{e}_2 = [0, -1/\sqrt{2}, 1/\sqrt{2}]$.

Problem 2.61 Plot the image on the s_1 - s_2 plane of the tetrahedron whose vertices are located at $[0, 0, 0]$, $[0, 1, 0]$, $[0, 1/2, \sqrt{3}/2]$ and $[\sqrt{3}/6, \sqrt{6}/3, 1/2]$ (Use the same values as before: $\mathbf{p} = [-10, 0, 0]$, $\mathbf{q} = [10, 0, 0]$, $\mathbf{e}_1 = [0, 1, 0]$, $\mathbf{e}_2 = [0, 0, 1]$ and $\mathbf{e}_3 = [-1, 0, 0]$.)

Problem 2.62 Suppose that that points \mathbf{x} lie on the line $\mathbf{x} = \mathbf{x}_0 + t\mathbf{v}$. Show that corresponding planar points $[s_1, s_2]$ also lie on a line. (Hint: show that there are numbers a , b , c that do not depend on t , so that $as_1 + bs_2 = c$ for every t .)

Problem 2.63 Consider a different drawing procedure where the point \mathbf{x} maps to the point on the view plane given by the intersection of the plane with the line through \mathbf{x} parallel to \mathbf{e}_3 . Find a formula for the s_1 and s_2 co-ordinates of \mathbf{x} .

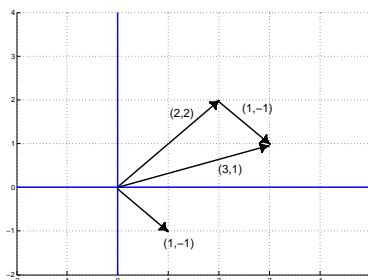


Figure 2.25: Solution to problem 2.1

2.8 Solutions to Chapter Problems

Solution 2.1 See Figure 2.25.

Solution 2.2

- (i) A straight line passing through the origin in the direction of \mathbf{a} .
- (ii) A ray (half line) passing through the origin in the direction of \mathbf{a} .
- (iii) A straight line parallel to \mathbf{a} passing through \mathbf{b} .
- (iv) If \mathbf{a} and \mathbf{b} do not lie on the same line: a plane passing through the origin parallel to \mathbf{a} and \mathbf{b} (in two dimensions, this is the whole plane). If \mathbf{a} and \mathbf{b} lie on the same line: a straight line passing through the origin in the direction of \mathbf{a} (and \mathbf{b}).
- (v) If \mathbf{a} and \mathbf{b} both lie on the same line: a line passing through \mathbf{c} parallel to \mathbf{a} (and \mathbf{b}). If \mathbf{a} and \mathbf{b} do not lie on the same line: a plane passing through \mathbf{c} parallel to \mathbf{a} and \mathbf{b} (in two dimensions, this is the whole plane).

Solution 2.3 $\mathbf{a} - \mathbf{b}$ is the vector that when added to \mathbf{b} gives \mathbf{a} . So if we draw it with its tail at \mathbf{b} then its head is at \mathbf{a} . Similarly $\mathbf{b} - \mathbf{a}$ when drawn with its tail at \mathbf{a} has its head at \mathbf{b} .

Solution 2.4 To find the midpoint we add half of the vector with its tail at \mathbf{a} and head at \mathbf{b} to \mathbf{a} . So the midpoint is $\mathbf{a} + (1/2)(\mathbf{b} - \mathbf{a}) = (1/2)(\mathbf{a} + \mathbf{b})$. Similarly, the point one third of the way from \mathbf{a} to \mathbf{b} is $\mathbf{a} + (1/3)(\mathbf{b} - \mathbf{a}) = (2/3)\mathbf{a} + (1/3)\mathbf{b}$.

Solution 2.5 The line segment is given by the vectors $\mathbf{a} + t(\mathbf{b} - \mathbf{a})$ as t varies between 0 and 1.

Solution 2.6 $\mathbf{a} = (1, 2)$; $\mathbf{b} = (1, -2)$

(a) $\mathbf{a} + \mathbf{b} = (2, 0)$

(b) $2\mathbf{a} = (2, 4)$

(c) $\mathbf{a} - \mathbf{b} = (0, 4)$

(d) $\mathbf{a} \cdot \mathbf{b} = 1 - 4 = -3$.

(e) $\|\mathbf{b}\| = \sqrt{1 + 4} = \sqrt{5}$.

Solution 2.7 The radius of the circle is

$$\|(2, 5) - (3, 3)\| = \|(-1, 2)\| = \sqrt{1 + 4} = \sqrt{5}.$$

Thus using the standard equation for a circle

$$(x - 2)^2 + (y - 5)^2 = (\sqrt{5})^2 = 5.$$

Solution 2.8 The equation is $\|\mathbf{x} - \mathbf{a}\| = r$, or

$$\sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2 + (x_3 - a_3)^2} = r$$

or

$$(x_1 - a_1)^2 + (x_2 - a_2)^2 + (x_3 - a_3)^2 = r^2$$

Solution 2.9 The midpoint of the sphere is the midpoint of $[2, 1, 4]$ and $[4, 3, 10]$, i.e., $(1/2)([2, 1, 4] + [4, 3, 10]) = [3, 2, 7]$. The radius is half the distance between $[2, 1, 4]$ and $[4, 3, 10]$, i.e.,

$$(1/2)\sqrt{(2 - 4)^2 + (1 - 3)^2 + (4 - 10)^2} = \sqrt{11}. \text{ Thus the equation is}$$

$$(x_1 - 3)^2 + (x_2 - 2)^2 + (x_3 - 7)^2 = 11.$$

Solution 2.10

(a) $\mathbf{a} \cdot \mathbf{b} = -2 + 6 = 4$, $\theta = \arccos(4/(\sqrt{5}\sqrt{13})) = 1.05 \dots$

(b) $\mathbf{a} \cdot \mathbf{b} = 1$, $\theta = 1.249 \dots$

(c) $\mathbf{a} \cdot \mathbf{b} = 4$, $\theta = 0$

(d) $\mathbf{a} \cdot \mathbf{b} = 2$, $\theta = 1.079 \dots \dots$

(e) $\mathbf{a} \cdot \mathbf{b} = 0$, $\theta = \pi/2 \dots$

Solution 2.11 Consider the dot product to determine angles between vectors, not the dot product (which only applies to 3D and cannot distinguish angles θ in the range $(0, \pi/2)$ from those in $(\pi/2, \pi)$, as discussed in the solution to problem 2.10 in the online notes).

$$\begin{aligned} \text{(a) } \mathbf{a} &= (1, 1, 1), & \mathbf{b} &= (3, 1, -2) \\ \mathbf{a} \cdot \mathbf{b} &= 3 + 1 - 2 = 2 \\ \|\mathbf{a}\| &= \sqrt{3}, & \|\mathbf{b}\| &= \sqrt{3^2 + 1^2 + (-2)^2} = \sqrt{14}. \\ \cos \theta &= \frac{2}{\sqrt{3}\sqrt{14}}, & \theta &= \cos^{-1} \frac{2}{\sqrt{3}\sqrt{14}} \simeq 1.26 \text{ radians, or } \simeq 72.0^\circ \end{aligned}$$

$$\begin{aligned} \text{(b) Using } \mathbf{a} \cdot \mathbf{b} \text{ and } \|\mathbf{a}\| \text{ above,} \\ \text{proj}_{\mathbf{a}} \mathbf{b} &= \frac{2}{3}(1, 1, 1) = \left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right). \text{ Remember, } \text{proj}_{\mathbf{a}} \mathbf{b} \text{ should be in the direction of } \mathbf{a}. \end{aligned}$$

Solution 2.12

$$\begin{aligned} \mathbf{a} &= (1, 4, 0) & \|\mathbf{a}\| &= \sqrt{17} \\ \mathbf{b} &= (2, -1, 5) & \|\mathbf{b}\| &= \sqrt{4 + 1 + 25} = \sqrt{30} \end{aligned}$$

$$\text{(a) } \cos \theta = \frac{\mathbf{b} \cdot \mathbf{a}}{\|\mathbf{a}\| \|\mathbf{b}\|} = \frac{2-4}{\sqrt{17}\sqrt{30}} = \frac{-2}{\sqrt{510}} \text{ so}$$

$$\theta = \cos^{-1} \left(\frac{-2}{\sqrt{510}} \right) \approx 1.66 \text{ in radians or } \approx 95^\circ$$

$$\text{(b) } \text{proj}_{\mathbf{a}} \mathbf{b} = \frac{\mathbf{b} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} = \frac{-2}{17}(1, 4, 0) = (-2/17, -8/17, 0).$$

Solution 2.13 The dot product is zero if $-1 + 2 + s = 0$, i.e., if $s = -1$.

Solution 2.14 Let $\mathbf{a} = [1, 2, 3]$, $\mathbf{b} = [4, 0, 5]$ and $\mathbf{c} = [3, 4, 6]$. The sides of the triangle are in the directions of $\mathbf{a} - \mathbf{b} = [-3, 2, -2]$, $\mathbf{a} - \mathbf{c} = [-2, -2, -3]$, and $\mathbf{b} - \mathbf{c} = [1, -4, -1]$. To see if there are any right angles we compute $(\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{c}) = [-3, 2, -2] \cdot [-2, -2, -3] = 8$, $(\mathbf{a} - \mathbf{b}) \cdot (\mathbf{b} - \mathbf{c}) = [-3, 2, -2] \cdot [1, -4, -1] = -9$ and $(\mathbf{a} - \mathbf{c}) \cdot (\mathbf{b} - \mathbf{c}) = [-2, -2, -3] \cdot [1, -4, -1] = 9$. Since none of these are zero, there are no right angles.

Solution 2.15 If $[c_1, 1, c_2] = s[2, -2, 3]$, with s a scalar multiple, then $s = -\frac{1}{2}$ (from the values in the second component), hence:
 $c_1 = -1$, $c_2 = -3/2$.

Solution 2.16 The problem becomes much simpler if we choose the y axis to run along the runway and the x axis to run perpendicular to the runway. Replace the wind and plane velocities by vectors \mathbf{w} and \mathbf{p} as on the diagram in Figure 2.26.

The wind vector has components $[-10 \cos(20^\circ), -10 \sin(20^\circ)]$ while the plane velocity has components $[70 \sin(\theta), 70 \cos(\theta)]$. We want the x component to be zero hence we need $10 \cos(20^\circ) = 70 \sin(\theta)$. This gives $\theta = 7.7^\circ$, i.e., we get the heading of the plane we have to add 7.7° to the runway direction. Thus the

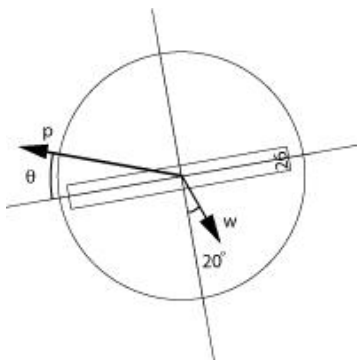


Figure 2.26: Coordinate system for Solution 2.16.

plane's heading is 267.7 . The groundspeed is the magnitude of the velocity, in this case simply the y component given by $70 \cos(7.7^\circ) - 10 \sin(20^\circ) = 66$ knots.

Solution 2.17 If we use the co-ordinate system where the x axis is horizontal and the y axis is vertical, then the force of gravity is $\mathbf{F}_g = [0, -mg]$. The unit vector in the direction along the shaft is $\mathbf{p} = [-\sin(\theta), \cos(\theta)]$. The force \mathbf{F}_s exerted by the shaft is in the direction of the shaft, hence a multiple of \mathbf{p} . So $\mathbf{F}_s = t\mathbf{p}$ for some number t . To find t we use that the component of the total force in the direction of the shaft must be zero. Thus $\text{proj}_{\mathbf{p}} \mathbf{F}_g + \mathbf{F}_s = 0$, i.e. $\mathbf{p} \cdot \mathbf{F}_g + t = 0$, i.e., $t = mg \cos(\theta)$. Thus the total force is $\mathbf{F}_g + \mathbf{F}_s = [0, -mg] + mg \cos(\theta)[- \sin(\theta), \cos(\theta)] = mg[-\cos(\theta) \sin(\theta), -1 + \cos^2(\theta)]$. Note that this is orthogonal to \mathbf{p} , as it must be.

If we use co-ordinates where the y axis runs along the shaft of the pendulum, and the x axis perpendicular to it, then $\mathbf{F}_g = [-mg \sin(\theta), -mg \cos(\theta)]$ while $\mathbf{F}_s = [0, t]$ for some value of t . In this case it is simply the y component that must be zero, so $t = mg \cos(\theta)$ and $\mathbf{F}_g + \mathbf{F}_s = [-mg \sin(\theta), 0]$

Solution 2.18

$$\sqrt{a(1)^2 + a(2)^2}$$

Solution 2.19 $[1, 2, 3] \times [4, 5, 6] = [-3, 6, -3]$

Solution 2.20

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

Since the determinant is zero, the volume of the parallelepiped generated by the row vectors is zero. This implies that the vectors lie on the same plane.

Solution 2.21 $\mathbf{a} \cdot (\mathbf{a} \times \mathbf{b}) = [a_1, a_2, a_3] \cdot [a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1] = a_1a_2b_3 - a_1a_3b_2 + a_2a_3b_1 - a_2a_1b_3 + a_3a_1b_2 - a_3a_2b_1 = 0$. Now $\mathbf{b} \cdot (\mathbf{a} \times \mathbf{b}) = -\mathbf{b} \cdot (\mathbf{b} \times \mathbf{a}) = 0$ by the previous calculation.

Solution 2.22

$$\begin{aligned} \text{(a)} \quad & ((1, 4, -1) \cdot (2, 1, 3))((2, 1, 4) \times (1, 4, 9)) \\ &= (2 + 4 - 3) \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 2 & 1 & 4 \\ 1 & 4 & 9 \end{vmatrix} \\ &= 3(\hat{i}(9 - 16) + \hat{j}(4 - 18) + \hat{k}(8 - 1)) \\ &= 3(-7, -14, 7) = (-21, -42, 21) \end{aligned}$$

$$\begin{aligned} \text{(b)} \quad & (7, 1, 0) \cdot ((2, 0, -1) \times (1, 4, 3)) \\ &= (7, 1, 0) \cdot \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 2 & 0 & -1 \\ 1 & 4 & 3 \end{vmatrix} \\ &= (7, 1, 0) \cdot (\hat{i}(0 + 4) + \hat{j}(-1 - 6) + \hat{k}(8 - 0)) \\ &= (7, 1, 0) \cdot (4, -7, 8) = 28 - 7 + 0 = 21 \end{aligned}$$

$$\begin{aligned} \text{(c)} \quad & (\mathbf{a} \times \mathbf{b}) \times (\mathbf{b} \times \mathbf{a}) \\ &= (\mathbf{a} \times \mathbf{b}) \times (-\mathbf{a} \times \mathbf{b}) = -(\mathbf{a} \times \mathbf{b}) \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0} \end{aligned}$$

because $\mathbf{c} \times \mathbf{c} = \mathbf{0}$ for any vector \mathbf{c} .

Solution 2.23 Here we are assuming that θ lies in $[0, \pi]$ so that $\sin(\theta)$ is positive. The quantity $\|\mathbf{a}\|$ is the length of the base of the parallelogram while $\|\mathbf{b}\| \sin(\theta)$ is the height. The area of a parallelogram is the product of these.

Solution 2.24 Since $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$ it is never true that $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$, unless $\mathbf{a} \times \mathbf{b} = \mathbf{0}$. For the other example, just try \mathbf{i} , \mathbf{j} and \mathbf{k} . We have $\mathbf{i} \times \mathbf{j} = \mathbf{k}$ so $\mathbf{i} \times (\mathbf{i} \times \mathbf{j}) = \mathbf{i} \times \mathbf{k} = -\mathbf{j}$ On the other hand $(\mathbf{i} \times \mathbf{i}) \times \mathbf{j} = \mathbf{0} \times \mathbf{j} = \mathbf{0}$

Solution 2.25

```
a=rand(3,1);
b=rand(3,1);
c=rand(3,1);
cross(a,b) - cross(b,a)
cross(a,cross(b,c)) - cross(cross(a,b),c)
```

This actually constitutes a proof, since we found a counter example. We don't have to show that it's always the case, we only had to show that in general the two assertions are not true.

Solution 2.26 This is just a long calculation. One way to break it up is to consider $\mathbf{a} = \mathbf{i}$, \mathbf{j} and \mathbf{k} separately. Suppose we can prove it for these special cases. Then write a general \mathbf{a} as $a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$. Then $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = a_1\mathbf{i} \times (\mathbf{b} \times \mathbf{c}) + a_2\mathbf{j} \times (\mathbf{b} \times \mathbf{c}) + a_3\mathbf{k} \times (\mathbf{b} \times \mathbf{c})$. Assuming for the moment we know that the special cases hold, then this equals $a_1((\mathbf{i} \cdot \mathbf{c})\mathbf{b} - (\mathbf{i} \cdot \mathbf{b})\mathbf{c}) + a_2((\mathbf{j} \cdot \mathbf{c})\mathbf{b} - (\mathbf{j} \cdot \mathbf{b})\mathbf{c}) + a_3((\mathbf{k} \cdot \mathbf{c})\mathbf{b} - (\mathbf{k} \cdot \mathbf{b})\mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$. Now we still have to prove the special cases. For example if $\mathbf{a} = \mathbf{i}$ we have $\mathbf{i} \times (\mathbf{b} \times \mathbf{c}) =$

$$\det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 0 \\ b_2c_3 - b_3c_2 & b_3c_1 - b_1c_3 & b_1c_2 - b_2c_1 \end{bmatrix} = -(b_1c_2 - b_2c_1)\mathbf{j} + (b_3c_1 - b_1c_3)\mathbf{k} = [0, -b_1c_2 + b_2c_1, b_3c_1 - b_1c_3]$$

On the other hand $(\mathbf{i} \cdot \mathbf{c})\mathbf{b} - (\mathbf{i} \cdot \mathbf{b})\mathbf{c} = c_1[b_1, b_2, b_3] - b_1[c_1, c_2, c_3] = [0, c_1b_2 - b_1c_2, c_1b_3 - b_1c_3]$. The other two are similar.

Solution 2.27

```
a=rand(3,1);
b=rand(3,1);
c=rand(3,1);
cross(a,cross(b,c)) - (dot(a,c)*b - dot(a,b)*c)
```

Solution 2.28 $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = ((\mathbf{a} \times \mathbf{b}) \times \mathbf{c}) \cdot \mathbf{d}$ (by property 5) $= -(\mathbf{c} \times (\mathbf{a} \times \mathbf{b})) \cdot \mathbf{d}$ (by property 1) $= -((\mathbf{c} \cdot \mathbf{b})\mathbf{a} + (\mathbf{c} \cdot \mathbf{a})\mathbf{b}) \cdot \mathbf{d} = -(\mathbf{c} \cdot \mathbf{b})(\mathbf{a} \cdot \mathbf{d}) + (\mathbf{c} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{d})$

Solution 2.29

- (a) Let $\mathbf{v} = \mathbf{a} \times (\mathbf{a} \times \mathbf{b})$. Since $\mathbf{a} \times \mathbf{b}$ is orthogonal to the plane, \mathbf{v} must lie in the plane of the paper. \mathbf{v} is also orthogonal to \mathbf{a} . Checking orientations with the right hand rule, there are only two sketches (up to rotation and resizing the vectors) shown in Figure 2.27.
- (b) $\mathbf{a} \times (\mathbf{a} \times \mathbf{b}) = (\mathbf{b} \cdot \mathbf{a})\mathbf{a} - (\mathbf{a} \cdot \mathbf{a})\mathbf{b}$ by property 2 in the notes.

$$= \|\mathbf{a}\|^2 \left(\frac{\mathbf{b} \cdot \mathbf{a}}{\|\mathbf{a}\|^2} \mathbf{a} - \mathbf{b} \right) = \|\mathbf{a}\|^2 (\text{proj}_{\mathbf{a}} \mathbf{b} - \mathbf{b}).$$

Solution 2.30 Actually there is no one right answer to this problem. But in two dimensions the analog of the cross product could be an operation on a single vector say $\times(\mathbf{a})$ with the property that $\times(\mathbf{a}) \cdot \mathbf{b} = \det \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$. This implies that $\times(\mathbf{a}) = [a_2, -a_1]$. Note that $\times(\mathbf{a})$ is orthogonal to \mathbf{a} . The answer in 4 dimensions is a little more obscure. In this case the cross product could be an operation on three vectors that produces a vector that is orthogonal to all of them. One way of producing such a vector would be to do the formal calculation of a 4×4 determinant, where the first row contains the four unit vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and \mathbf{e}_4 , while the other three rows contain the entries of \mathbf{a}, \mathbf{b} and \mathbf{c} . Do you see why? (There is another possible, perhaps better answer, involving a different type of product called the wedge product.)

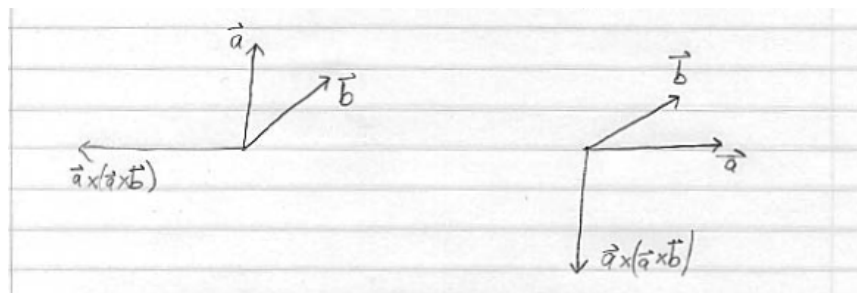


Figure 2.27: Vector sketch for the solution to 2.29.

Solution 2.31 Let \mathbf{q} be a point on the line. Then the projection of \mathbf{x} in the direction of \mathbf{b} is $\frac{\mathbf{x} \cdot \mathbf{b}}{\|\mathbf{b}\|^2}$. The set of all points whose projection in the direction of \mathbf{b} are the same as the projection of the point \mathbf{q} in the direction of \mathbf{b} is therefore the set of points \mathbf{x} such that $\frac{\mathbf{x} \cdot \mathbf{b}}{\|\mathbf{b}\|^2} = \frac{\mathbf{q} \cdot \mathbf{b}}{\|\mathbf{b}\|^2}$. Multiplying both sides by $\|\mathbf{b}\|^2$ gives $\mathbf{x} \cdot \mathbf{b} = \mathbf{q} \cdot \mathbf{b}$.

Solution 2.32 The line is in the direction of $\mathbf{a} = [7, 5] - [0, 1] = [7, 4]$ and is therefore perpendicular to $\mathbf{b} = [4, -7]$. $\mathbf{q} = [0, 1]$ is a point on the line. So the parametric form for the line is $[0, 1] + s[7, 4]$ and the equation form is $4x_1 - 7x_2 = -7$. So five points on the line are (picking $s = -2, 1, 0, 1, 2$) $[-14, -7], [-7, -3], [0, 1], [7, 5], [14, 9]$. To check whether $[1012/3, 1069/21]$ lies on the line, plug it into the equation: $4 \cdot 1012/3 - 7 \cdot 1069/21 \neq -7$ so the point does not lie on the line.

Solution 2.33 We can take $\mathbf{q} = [1, 1]$ and $\mathbf{b} = [2, 1]$ so the equation is $2x_1 + x_2 = 3$.

Solution 2.34 We have to find a point on the line. Try for one of the form $[t, 0]$. This is on the line if $t - 3 \cdot 0 = 5$. So $[5, 0]$ is on the line. The vector orthogonal to the line is $[1, -3]$ so a vector parallel is $[3, 1]$. Thus the parametric form is $[5, 0] + s[3, 1]$.

Solution 2.35 Pick a point \mathbf{q} on the line and let \mathbf{b} be orthogonal to the line. Then the distance from a point \mathbf{x} to the line is the length of the projection of $\mathbf{x} - \mathbf{q}$ onto \mathbf{b} as shown in Figure 2.28. In this case we can take $\mathbf{b} = [3, -4]$, $\mathbf{q} = [0, 1]$ and $\mathbf{x} = [-2, 3]$. The projection is $((\mathbf{x} - \mathbf{q}) \cdot \mathbf{b} / \|\mathbf{b}\|^2) \mathbf{b}$. This equals $([-2, 2] \cdot [3, -4] / 25) [3, -4] = (-14/25) [3, 4]$. The length of this vector is $14/5$.

Solution 2.36

- the normal direction to the plane is $(1, -1, 2)$
- the point $(7, 0, 0)$ is on the plane (look for a solution with $y = z = 0$ that is the intersection of the plane with the line of the x-axis).

Solution 2.37 The midpoint between \mathbf{b} and \mathbf{c} is $(\mathbf{b} + \mathbf{c})/2$. So the vector from \mathbf{a} to this midpoint is $(\mathbf{b} + \mathbf{c})/2 - \mathbf{a}$ and thus the parametric form for the median

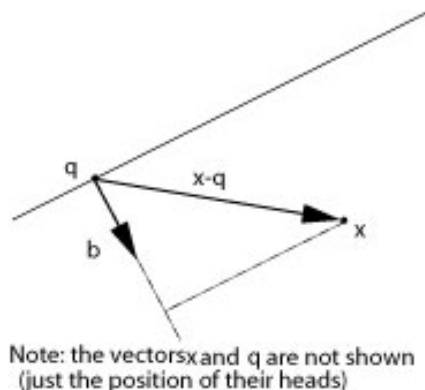


Figure 2.28: The projection used in Solution 2.35.

is $\mathbf{a} + t((\mathbf{b} + \mathbf{c})/2 - \mathbf{a})$. Similarly the parametric forms for the other medians is $\mathbf{b} + s((\mathbf{a} + \mathbf{c})/2 - \mathbf{b})$ and $\mathbf{c} + r((\mathbf{a} + \mathbf{b})/2 - \mathbf{c})$. When $t = s = r = 2/3$ these lines meet at the common point $(\mathbf{a} + \mathbf{b} + \mathbf{c})/3$

Solution 2.38 The line is in direction $(0, 1, 3)$. Two vectors which are orthogonal to the line are $(1, 0, 0)$ and $(0, 3, -1)$, because $(1, 0, 0) \cdot (0, 1, 3) = 0$ and $(0, 3, -1) \cdot (0, 1, 3) = 0$. Note that there are other choices for these two vectors that lead to different correct answers below. There will be two equations of the form

$$\begin{aligned}x &= c_1 \\3y - z &= c_2\end{aligned}$$

Substituting the point $(2, 0, 4)$ that we know is on the line gives $c_1 = 2$ and $c_2 = 4$. Thus an equation form of the line is

$$\begin{aligned}x &= 2 \\3y - z &= 4\end{aligned}$$

Solution 2.39 To find an equation for the plane we need an orthogonal vector

$$(0, 1, -3) \times (-1, 2, 0) = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 0 & 1 & -3 \\ -1 & 2 & 0 \end{vmatrix} = \hat{i}(0+6) + \hat{j}(3-0) + \hat{k}(0+1) = (6, 3, 1).$$

So an equation for the plane is $6x + 3y + z = c$ and we know that $c = 0$ since the plane goes through the origin. Any point on the line has the form

$$(2, -1, 6) + s(1, -1, 0) = (2 + s, -1 - s, 6).$$

Substituting this into the equation for the plane gives

$$6(2 + s) + 3(-1 - s) + 6 = 0$$

$$\begin{aligned} 3s + 15 &= 0 \\ s &= -5 \end{aligned}$$

So the intersection point is

$$(2, -1, 6) - 5(1, -1, 0) = (2, -1, 6) + (-5, 5, 0) = (-3, 4, 6).$$

Note: it can be checked that $t = -2$ and $u = 3$ in the parametric form of the plane gives this point as well.

Solution 2.40 Points on the line satisfy

$$\begin{aligned} x &= 1 + t \\ y &= 2 \\ z &= 3 - t. \end{aligned}$$

To be on the plane, $x + 2y - z = 5$, or $(1 + t) + 2(2) - (3 - t) = 5$, or $2t + 2 = 5$, or $t = 3/2$.

Thus the intersection occurs at $(1, 2, 3) + (\frac{3}{2})(1, 0, -1) = (\frac{5}{2}, 2, \frac{3}{2})$.

Solution 2.41 Two vectors in the direction of the plane are $[1, 1, 0] - [1, 0, 1] = [0, 1, -1]$ and $[0, 1, 1] - [1, 0, 1] = [-1, 1, 0]$. To find a normal vector take the cross product $[0, 1, -1] \times [-1, 1, 0] = [1, 1, 1]$. So the equation of the plane is $x_1 + x_2 + x_3 = 2$.

Solution 2.42 The centre of the sphere lies in the plane halfway between the planes $x_1 + x_2 + x_3 = 3$ and $x_1 + x_2 + x_3 = 9$. This is the plane $x_1 + x_2 + x_3 = 6$. So the centre satisfies the three equations

$$\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ 2x_1 - 3x_2 &= 0 \\ 3x_1 - x_3 &= 0 \end{aligned}$$

We will develop efficient techniques for solving such systems of equations. For now, we can just use brute force: The second equation says $x_2 = 2x_1$ and the third $x_3 = 3x_1$. Substituting this into the first equation yields $x_1 = 1$, so $x_2 = 2$ and $x_3 = 3$. Therefore the centre of the sphere is $[1, 2, 3]$. To compute the radius, we must find the distance between the planes $x_1 + x_2 + x_3 = 3$ and $x_1 + x_2 + x_3 = 6$. If we divide the equations by $\|[1, 1, 1]\| = \sqrt{3}$, the first can be interpreted as vectors whose projection onto the direction $\|[1, 1, 1]\|$ has length $3/\sqrt{3} = \sqrt{3}$, the second can be interpreted as vectors whose projection onto the direction $\|[1, 1, 1]\|$ has length $6/\sqrt{3} = 2\sqrt{3}$. The radius is $\sqrt{3}$ so the equation is $(x_1 - 1)^2 + (x_2 - 2)^2 + (x_3 - 3)^2 = 3$.

Solution 2.43

$$\begin{aligned} x + y + z = 2 & \quad \text{normal } (1, 1, 1) = \vec{b}_1 \\ x - y + 2z = 7 & \quad \text{normal } (1, -1, 2) = \vec{b}_2 \end{aligned}$$

The line direction can be found as

$$\vec{a} = \vec{b}_1 \times \vec{b}_2 = \begin{bmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix} = (3, -1, -2).$$

We also need a point on the line. We can look for a point which has $z = 0$ (the point that is the intersection of the line and the x-y plane). We get the two by two system

$$x + y = 2, \quad x - y = 7$$

Solving the system yields that $x = \frac{9}{2}$, and $y = -\frac{5}{2}$.

So $(\frac{9}{2}, -\frac{5}{2}, 0) + s(3, -1, -2)$ is a parametric form of the line with parameter s .

Solution 2.44 Set $x_2 = 0$ and solve $2x_1 - x_3 = 0$ and $x_1 + 2x_3 = -5$ giving $x_1 = -1$ and $x_3 = -2$. This implies $[-1, 0, -2]$ is in the plane we are looking for. Set $x_3 = 0$ and solve $2x_1 + 3x_2 = 0$ and $x_1 - 4x_2 = -5$ giving $x_1 = -15/11$ and $x_2 = 10/11$. Hence $[-15/11, 10/11, 0]$ also lies on the plane. We now have three points on the plane. Two vectors in the direction of the plane are therefore $[-2, 0, 1] - [-1, 0, -2] = [-1, 0, 3]$ and $[-15/11, 10/11, 0] - [-1, 0, -2] = [-4/11, 10/11, 2]$. Thus a normal vector is $[-1, 0, 3] \times [-4/11, 10/11, 2] = [-30/11, 10/11, -10/11]$. This is parallel to $[-3, 1, -1]$. Hence the equation is $-3x_1 + x_2 - x_3 = 5$.

Solution 2.45 The three points are all on the same line. To see this, notice that the vectors $[2, 3, 4] - [1, 2, 3] = [1, 1, 1]$ and $[3, 4, 5] - [1, 2, 3] = [2, 2, 2]$ are parallel.

Solution 2.46 The distance of \mathbf{p} to the plane $\mathbf{b} \cdot \mathbf{x} = c$ is the length of the projection of $\mathbf{p} - \mathbf{q}$ onto \mathbf{b} , where \mathbf{q} is any point on the plane. A point on the plane is $[c/b_1, 0, 0]$ (unless $b_1 = 0$ in which case we choose either $[0, c/b_2, 0]$ or $[0, 0, c/b_3]$). The length of the projection is $(\mathbf{p} - \mathbf{q}) \cdot \mathbf{b} / \|\mathbf{b}\| = (\mathbf{b} \cdot \mathbf{p} - c) / \|\mathbf{b}\|$.

Solution 2.47 If the line is parallel to both planes, then it is orthogonal to both normal vectors. Therefore $[1, 1, 0] \times [1, -1, 2] = [2, -2, -2]$ is in the direction of the line. Therefore the parametric form of the line is $[2, -1, -1] + s[2, -2, -2]$. The two equations are $x_1 + x_2 = [1, 1, 0] \cdot [2, -1, -1] = 1$ and $x_1 - x_2 + 2 = [1, -1, 2] \cdot [2, -1, -1] = 1$.

Solution 2.48

```
x=-3:0.1:3
y = sqrt(9-x.^2)
plot(x,y, 'r')
hold on
plot(x,-y, 'r')
axis equal
```

Solution 2.49 Yes $[1, 1]$ and $[1, 0]$ do form a basis, since they don't lie on the same line. (By the way, there is some potential for confusion here. When I say here that $[1, 1]$ and $[1, 0]$ don't lie on the same line, I'm thinking of them as position vectors, drawn with their tails at the origin. It probably would be more clear to say that the vectors are not parallel.) Every vector in the plane can be written as a linear combination of these two. In particular $[0, 1] = [1, 1] - [1, 0]$

Solution 2.50 Note that \mathbf{a} and \mathbf{b} are linearly independent (they are not multiples of one another). So \mathbf{a} , \mathbf{b} and \mathbf{c} will be linearly independent and so a basis for \mathbb{R}^3 unless \mathbf{c} is a linear combination of \mathbf{a} and \mathbf{b} , that is it can be written as

$$\mathbf{c} = r\mathbf{a} + s\mathbf{b}$$

for some r and s .

Solution 2.51 The collection is a basis if the vectors are linearly independent. We consider $r\mathbf{a} + s\mathbf{b} + t\mathbf{c} = \mathbf{0}$ which leads to

$$\begin{aligned} r + s + t &= 0 \\ r + s &= 0 \\ r &= 0 \end{aligned}$$

Starting from the last equation and working up, we see that $r = 0$, $s = 0$, and $t = 0$ which shows that the vectors are linearly independent and therefore form a basis. The linear independence could also have been shown by considering the determinant

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = 1 \neq 0$$

Since the determinant is not zero the vectors do not lie on the same plane and so are linearly independent. To write $[1, 2, 3]$ as a linear combination

$$[1, 2, 3] = r[1, 1, 1] + s[1, 1, 0] + t[1, 0, 0]$$

match components working backwards to get successively $r = 3$, $s = -1$ and $t = -1$.

Solution 2.52 The vectors \mathbf{a} , \mathbf{b} , and \mathbf{c} are vectors in 3-D. Thus they form a basis if and only if they are linearly independent.

The three vectors are linearly independent if and only if the matrix A with rows given by the three vectors satisfies

$$\det(A) \neq 0.$$

Here

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

Solution 2.53 No, it is not possible for four vectors to be linearly independent in three dimensional space. To see this, first notice that if four vectors are linearly independent any subset of three vectors must also be linearly independent (If three vectors would lie on the same plane, we could find a non-trivial linear combination of those three equal to zero. Then by adding 0 times the left over vector, we would get a non-trivial linear combination of all four vectors equal

to zero, contradicting their independence.) But this means that those three vectors form a basis for three dimensional space. So the fourth vector must be expressible as a linear combination of the first three. This means the four vectors are not independent.

Solution 2.54 It is enough to check that $\det(A) \neq 0$. We have that

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 0 & 2 \\ 3 & 0 & 0 \end{bmatrix}, \quad \det(A) = 6 \neq 0,$$

so they form a basis. We want to find s_1 , s_2 , and s_3 such that

$$\begin{pmatrix} 12 \\ 2 \\ 4 \end{pmatrix} = s_1 \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + s_2 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + s_3 \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}.$$

We have the following system of equations:

$$\begin{aligned} 12 &= 2s_1 + s_2 + 3s_3 \\ 2 &= s_1 \\ 4 &= 3s_1 + 2s_2 \end{aligned}$$

Hence $s_1 = 2$, $s_2 = -1$, $s_3 = 3$.

Solution 2.55 If $\mathbf{x} = s_1\mathbf{a}_1 + s_2\mathbf{a}_2 + \cdots + s_n\mathbf{a}_n$ and $\mathbf{x} = t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \cdots + t_n\mathbf{a}_n$, then $\mathbf{0} = \mathbf{x} - \mathbf{x} = (s_1 - t_1)\mathbf{a}_1 + (s_2 - t_2)\mathbf{a}_2 + \cdots + (s_n - t_n)\mathbf{a}_n$. Since the vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ are linearly independent, the only way this can happen is $s_1 - t_1 = 0, \dots, s_n - t_n = 0$, or $s_1 = t_1, \dots, s_n = t_n$.

Solution 2.56 The values $\alpha = 0.23$, $\beta = 0.15$ provide a good approximation to the linear combination.

Solution 2.57 The axis is parallel to $[2, -3, 1] - [1, 1, -1] = [1, -4, 2]$. The unit vector in this direction is $1/\sqrt{21}[1, -4, 2]$ (Actually there is an ambiguity in the problem: the unit vector could also point in the opposite direction. This would change the sign of the answer.) So $\boldsymbol{\Omega} = (10/\sqrt{21})[1, -4, 2]$. The vector \mathbf{r} has its tail co-incident with the tail of $\boldsymbol{\Omega}$ and its head at $[1, 2, 3]$. So $\mathbf{r} = [1, 2, 3] - [1, 1, -1] = [0, 1, 4]$ and the velocity is $\mathbf{v} = \boldsymbol{\Omega} \times [0, 1, 4] = (10/\sqrt{21})[-18, -4, 1]$

Solution 2.58 Some preliminary discussion:

- (i) it does not matter where the centre of the rod is along L, we can take it to be at the origin for convenience.
- (ii) the fastest speed is at the rod tip, and this does not change in time, so
- (iii) it is OK to evaluate the speed at the initial position.

$$\Omega = 3 \cdot 2\pi = 6\pi \text{ radians/min.}$$

$$\hat{a} = \frac{1}{\sqrt{3}}(1, 1, 1) \text{ unit vector in the direction of the axis of rotation.}$$

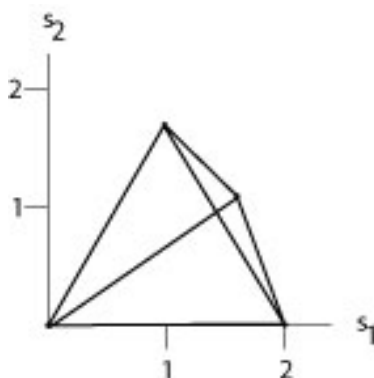


Figure 2.29: Mapped tetrahedron of Solution 2.61.

$\vec{v} = \Omega \hat{a} \times \vec{r} = \frac{6\pi}{\sqrt{3}}(1, 1, 1) \times (1, 0, 0)$, with \vec{r} the initial position of the tip.
Now evaluate the cross product:

$$\begin{bmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} = (0, 1, -1),$$

so $\vec{v} = \frac{6\pi}{\sqrt{3}}(0, 1, -1)$, and hence the speed is $\|\vec{v}\| = 2\pi\sqrt{6}\text{m/min}$

Solution 2.59 We have $[x(t), y(t), 0] = r[\cos(\theta(t)), \sin(\theta(t)), 0]$ so the velocity is

$$\begin{aligned} [\dot{x}(t), \dot{y}(t), 0] &= r[-\sin(\theta(t))\dot{\theta}(t), \cos(\theta(t))\dot{\theta}(t), 0] \\ &= r\dot{\theta}(t)[- \sin(\theta(t)), \cos(\theta(t)), 0] \end{aligned}$$

On the other hand $\boldsymbol{\Omega} = [0, 0, \dot{\theta}(t)]$ so that $\boldsymbol{\Omega} \times r[\cos(\theta(t)), \sin(\theta(t)), 0]$ gives the same answer.

Solution 2.60 We have $\mathbf{q} - \mathbf{p} = [20, 0, 0]$, $\mathbf{x} - \mathbf{p} = [11, 3, 2]$. Thus $\mathbf{e}_1 \cdot (\mathbf{q} - \mathbf{p}) = \mathbf{e}_2 \cdot (\mathbf{q} - \mathbf{p}) = 0$, $\mathbf{e}_3 \cdot (\mathbf{q} - \mathbf{p}) = -20$, $\mathbf{e}_1 \cdot (\mathbf{x} - \mathbf{p}) = 5\sqrt{2}/2$, $\mathbf{e}_2 \cdot (\mathbf{x} - \mathbf{p}) = \sqrt{2}/2$, $\mathbf{e}_3 \cdot (\mathbf{x} - \mathbf{p}) = -11$. So $s_1 = 50\sqrt{2}/11$ and $s_2 = 10\sqrt{2}/11$.

Solution 2.61 The image of the vertices are $[0, 0]$, $[2, 0]$, $[1, \sqrt{3}]$ and some more horrible expression which is approximately $[1.587\dots, 0.9719\dots]$. So the image on the screen looks something like what is shown in Figure 2.29. Notice that the first three points— $[0, 0]$, $[2, 0]$ and $[1, \sqrt{3}]$ —are simply double the x_2 and x_3 co-ordinates of the original points in space. Can you explain this geometrically? (I know these are supposed to be answers, not questions, but still ...) (By the way, I actually intended the last point in the question to be $[\sqrt{6}/3, 1/2, \sqrt{3}/6]$. Can you say why?)

Solution 2.62 We can write

$$[s_1, s_2] = \frac{1}{\mathbf{e}_3 \cdot (\mathbf{x}_0 - \mathbf{p}) + t\mathbf{e}_3 \cdot \mathbf{v}} [\mathbf{e}_1 \cdot (\mathbf{x}_0 - \mathbf{p}) + t\mathbf{e}_1 \cdot \mathbf{v}, \mathbf{e}_2 \cdot (\mathbf{x}_0 - \mathbf{p}) + t\mathbf{e}_2 \cdot \mathbf{v}]$$

This means that $as_1 + bs_2 = c$ can be rewritten

$$a(\mathbf{e}_1 \cdot (\mathbf{x}_0 - \mathbf{p}) + t\mathbf{e}_1 \cdot \mathbf{v}) + b(\mathbf{e}_2 \cdot (\mathbf{x}_0 - \mathbf{p}) + t\mathbf{e}_2 \cdot \mathbf{v}) - c(\mathbf{e}_3 \cdot (\mathbf{x}_0 - \mathbf{p}) + t\mathbf{e}_3 \cdot \mathbf{v}) = 0$$

This holds for every t if both these equations hold:

$$\begin{array}{rcccc} a(\mathbf{e}_1 \cdot (\mathbf{x}_0 - \mathbf{p})) & +b(\mathbf{e}_2 \cdot (\mathbf{x}_0 - \mathbf{p})) & -c(\mathbf{e}_3 \cdot (\mathbf{x}_0 - \mathbf{p})) & = & 0 \\ a\mathbf{e}_1 \cdot \mathbf{v} & +b\mathbf{e}_2 \cdot \mathbf{v} & -c\mathbf{e}_3 \cdot \mathbf{v} & = & 0 \end{array}$$

This is a system of two equations in three unknowns, a , b , and c . Such a system always has a non trivial solution (as we will see in the following chapter.)

Solution 2.63 In this case we want $\mathbf{y} - \mathbf{x}$ to point in the same direction as \mathbf{e}_3 so $\mathbf{y} - \mathbf{x} = \lambda\mathbf{e}_3$. On the other hand \mathbf{y} lies on the plane of the screen, so $\mathbf{y} = \mathbf{q} + s_1\mathbf{e}_1 + s_2\mathbf{e}_2$. Therefore

$$\mathbf{q} - \mathbf{x} + s_1\mathbf{e}_1 + s_2\mathbf{e}_2 = \lambda\mathbf{e}_3.$$

Now to find s_1 and s_2 take dot products with \mathbf{e}_1 and \mathbf{e}_2 . This yields $s_1 = -\mathbf{e}_1 \cdot (\mathbf{q} - \mathbf{x})$ and $s_2 = -\mathbf{e}_2 \cdot (\mathbf{q} - \mathbf{x})$.

Chapter 3

Solving Linear Systems

3.1 Linear Systems

3.1.1 General Form of Linear Systems

So far, we have seen systems of linear equations as the equations that describe points, lines and planes. However, linear systems of equations show up in many other ways in engineering problems. We will solve linear systems to find the behaviour of an electrical circuit. Other examples would be the calculation of equilibrium temperature distributions or electric fields. Such examples often involve the discretization of a continuous function. In other words, a continuous function like the temperature distribution in a body (which has a value for each of the infinitely many points in the body) will be replaced by a list of temperatures at a large but finite number n of closely spaced points. This gives rise to a system of linear equations in n unknowns, where n can be in the tens of thousands, or higher. Therefore, we want to develop a technique to solve systems of linear equations in n unknowns when n is large.

The most general form of a linear system of equations is

$$\begin{array}{cccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + \cdots & + & a_{1,n}x_n & = & c_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + \cdots & + & a_{2,n}x_n & = & c_2 \\ \vdots & & \vdots & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + \cdots & + & a_{m,n}x_n & = & c_m \end{array}$$

Here the numbers $a_{i,j}$ and c_j are known, and the goal is to find all values of x_1, \dots, x_n that satisfy all the equations.

3.1.2 Solving Linear Systems by Substitution

Let us start with an example, which we will solve using the method of *substitution*.

Example 3.1 Consider the system of equations

$$\begin{aligned}x_1 + x_2 + x_3 &= 6 \\x_1 - x_2 + x_3 &= 0 \\2x_1 + x_2 - 8x_3 &= -11\end{aligned}$$

One could try to proceed as follows. Solve the first equations for, say, x_3 . This gives

$$x_3 = 6 - x_1 - x_2.$$

Now substitute this value for x_3 into the second and third equations. This gives

$$\begin{aligned}x_1 - x_2 + (6 - x_1 - x_2) &= 0 \\2x_1 + x_2 - 8(6 - x_1 - x_2) &= -11\end{aligned}$$

or

$$\begin{aligned}-2x_2 &= -6 \\10x_1 + 9x_2 &= 37\end{aligned}$$

Now solve the first of these equations for x_2 and substitute into the last equation. This gives $x_2 = 3$ and $x_1 = 1$. Finally we can go back and calculate $x_3 = 6 - 1 - 3 = 2$.

Although this procedure works fine for $n = 2$ or even $n = 3$, it rapidly becomes unwieldy for larger values of n . We will now introduce a technique called Gaussian elimination that works well for large n and can be easily implemented on a computer.

We have already observed that there may be many systems of equations with the same solution. When there are only two unknowns, this amounts to saying that different pairs of lines may intersect in the same point. Gaussian elimination is based on the following idea. We introduce three elementary row operations. These operations change the system of the equations into another system with exactly the same the set of solutions. We then apply these elementary row operations in a systematic way to change the system of equations into a system that is easily solved.

3.1.3 Elementary row (equation) operations

The first elementary row operation is

1. Multiplication of a row (equation) by a non-zero number

For example, if we multiply the first equation in the system of Example 3.1 above by 3, we end up with

$$\begin{aligned}3x_1 + 3x_2 + 3x_3 &= 18 \\x_1 - x_2 + x_3 &= 0 \\2x_1 + x_2 - 8x_3 &= -11\end{aligned}$$

This new system of equations has exactly the same solutions as the original system, because we can undo the elementary row operation simply by dividing the first equation by 3. Thus the values x_1, x_2, x_3 solve this system if and only if they solve the original system. (Notice that this would not be true if we multiplied by zero. In that case we could not undo the operation, and the new system of equations could well have more solutions than the original system.) Any row operations we do can be undone by other row operations, and the set of solutions of the linear system remain unchanged.

The second elementary row operation is

2. *Adding a multiple of one row (equation) to another row*

For example, if we added 2 times the first row to the second row in our example we would obtain the system

$$\begin{array}{rclcl} x_1 & + & x_2 & + & x_3 & = & 6 \\ 3x_1 & + & x_2 & + & 3x_3 & = & 12 \\ 2x_1 & + & x_2 & - & 8x_3 & = & -11 \end{array}$$

Again, the new system of equations has exactly the same solutions as the original system, since we could undo this elementary row operation by subtracting 2 times the first row from the second row.

The third and final elementary row operation is

3. *Interchanging two rows (equations)*

For example, if we swapped the first and second equations in our original system we would end up with

$$\begin{array}{rclcl} x_1 & - & x_2 & + & x_3 & = & 0 \\ x_1 & + & x_2 & + & x_3 & = & 6 \\ 2x_1 & + & x_2 & - & 8x_3 & = & -11 \end{array}$$

This obviously doesn't change the solutions of the system since we have the same equalities.

3.1.4 Augmented Matrices

To save unnecessary writing, we now set up an streamlined notation for systems of linear equations. Notice that the only thing that distinguished one system of equations from another are the coefficients. So, as shorthand, we can write the system of equations of Example 3.1

$$\begin{array}{rclcl} x_1 & + & x_2 & + & x_3 & = & 3 \\ x_1 & - & x_2 & + & x_3 & = & 3 \\ 2x_1 & + & x_2 & - & 8x_3 & = & -4 \end{array}$$

simply as

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & -1 & 1 & 3 \\ 2 & 1 & -8 & -4 \end{array} \right]$$

This is called an augmented matrix. “Augmented” refers to the column to the right of the line that contains the information about the right side of each equation.

3.1.5 Problems

Problem 3.1 Express the system

$$\begin{aligned}x_1 - 2x_2 + 3x_3 &= 6 \\4x_1 - 5x_2 - 6x_3 &= 7 \\8x_1 + 9x_2 + 10x_3 &= 11\end{aligned}$$

as an augmented matrix.

Problem 3.2 Start with the system

$$\begin{aligned}x_1 + x_2 + x_3 &= 6 \\x_1 - x_2 + x_3 &= 0 \\2x_1 + x_2 - 8x_3 &= -11\end{aligned}$$

and perform the following sequence of row operations:

1. Subtract the first row from the second row
2. Subtract twice the first row from the third row
3. Multiply the second row by $-1/2$
4. Add the second row to the third row
5. Multiply the third row by $-1/10$

Solve the resulting system of equations by starting with the third equation, then the second and then the first.

Problem 3.3 (Matlab) Let's try to solve problem 3.2 using Matlab.

1. Start by generating an extended matrix to represent the system of equations:

$$A=[1, 1, 1, 6; 1, -1, 1, 0; 2, 1, -8, -11]$$

2. Subtract the first row from the second row:

$$A(2,:) = A(2,:) - A(1,:)$$

3. Subtract twice the first row from the third row:

$$A(3,:) = A(3,:) - 2*A(1,:)$$

4. Multiply the second row by $-1/2$:

$$A(2, :) = -A(2, :)/2$$

5. etc.

What other commands are necessary to arrive at the solution? Does it coincide with the paper and pen solution?

Problem 3.4 Start with the system

$$\begin{array}{rccccrcr} x_1 & - & x_2 & + & x_3 & = & 10 \\ x_1 & - & x_2 & - & x_3 & = & 6 \\ 6x_1 & + & 3x_2 & + & x_3 & = & 0 \end{array}$$

Perform the following sequence of row operations

1. Subtract the second row from the first row
2. Add the first row to the second row
3. Subtract the first row from the third row
4. Multiply the second row by 3
5. Add the second row to the third row

Show that the resulting system of equations can be easily solved and find the solution of the above system of equations.

Problem 3.5 Without using the calculator or computer, find the solution to the system

$$\begin{array}{rccr} 2x_1 & + & x_2 & = & 5 \\ 3x_1 & + & 5x_2 & = & -10 \end{array}$$

You can leave the solutions as fractions.

3.2 Gaussian Elimination

Recall that we want to use a sequence of elementary row operations to turn an arbitrary system of equations into an easily solved system of equations (with exactly the same solutions). What equations are easily solved? Well, the easiest possible equations to solve are ones of the form below, written as an augmented matrix:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -4 \end{array} \right]$$

If we translate from the shorthand back to the equations they represent, the first row says $x_1 = 3$, the second row says $x_2 = 3$ and the third row says $x_3 = -4$. In other words, we can just read off the values of x_1 , x_2 and x_3 in the rightmost column. The equations are already solved, and there is nothing left to do!

Slightly more work, but still easy to do, are upper triangular systems. These are systems where all the entries below the diagonal are equal to zero, as in

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & -4 & -8 \end{array} \right]$$

The reason these are easy to solve is that the equation represented by the j th row only involves the variables x_j, x_{j+1}, \dots, x_n . So if we start with the last equation (in the example $-4x_3 = -8$), we can solve it immediately for x_n (in the example $x_3 = 2$). Now we move up one equation. This equation only involves x_{n-1} and x_n , and we already know x_n . So we can solve it for x_{n-1} (in the example $-x_2 + x_3 = 3$ so $-x_2 + 2 = 3$ so $x_2 = -1$). We can continue in this way until all the x_n 's have been found. (In the example there is one more equation $x_1 + x_2 + x_3 = 3$ or $x_1 - 1 + 2 = 3$ or $x_1 = 2$.)

In practise (i.e., in a typical computer program) the procedure that is actually used is to apply a sequence of row operations to turn the system of equations into an upper triangular system. Then the equations are solved one by one, starting at the bottom and working up. This is the most efficient way to solve a system of equations. However, its sometimes convenient to apply row operations to bring the equation into the “completely solved” form. Then, you can just read off the solution from the last column.

Let us now do a bunch of examples to illustrate this procedure. I'll cook them up so that everything that possibly can go wrong, does go wrong. (I'll also cook them up so that the numbers come out looking nice. This will definitely not be the case in an example coming up in a real application!). Here is a shorthand for indicating which elementary row operation was done. The notation $3(1, :)$ means the first row was multiplied by the non-zero number 3. The notation $(2, :) = (2, :) - 4(5, :)$ means that 4 times the fifth row was subtracted from the second row. Finally, $(2, :) \leftrightarrow (3, :)$ means that the second and third row were interchanged.

Example 3.2 *Let us start with*

$$\left[\begin{array}{cccc|c} 1 & 2 & -2 & -7 & -29 \\ 1 & 2 & -1 & -5 & -18 \\ 0 & 3 & 0 & -3 & -6 \\ -1 & 4 & 1 & 1 & 14 \end{array} \right]$$

We are trying to put this matrix in upper triangular form. So we start by trying to produce zero entries in the first column under the top entry. We can do this by adding multiples of the first row to the other rows. So, the first move is to subtract the first row from the second row. The result is

$$\left[\begin{array}{cccc|c} 1 & 2 & -2 & -7 & -29 \\ 0 & 0 & 1 & 2 & 11 \\ 0 & 3 & 0 & -3 & -6 \\ -1 & 4 & 1 & 1 & 14 \end{array} \right] \quad (2, :) = (2, :) - (1, :)$$

The third row already has a zero in the first column, so there is nothing to do here. To put a zero in the fourth row we add the first row to the last row.

$$\left[\begin{array}{cccc|c} 1 & 2 & -2 & -7 & -29 \\ 0 & 0 & 1 & 2 & 11 \\ 0 & 3 & 0 & -3 & -6 \\ 0 & 6 & -1 & -6 & -15 \end{array} \right] \quad (4, :) = (4, :) + (1, :)$$

Now we shift our attention to the second column. We want to produce zeros below the diagonal. If we attempt to do this by adding multiples of the first row to other rows, we will destroy the zeros that we have already produced. So we try to use the second row. This is where we run into the first glitch. Since the entry in the second column of the second row is zero, adding a multiple of this row to the others won't have any effect on the numbers in the second column that we are trying to change. To remedy this we simply swap the second and third rows.

$$\left[\begin{array}{cccc|c} 1 & 2 & -2 & -7 & -29 \\ 0 & 3 & 0 & -3 & -6 \\ 0 & 0 & 1 & 2 & 11 \\ 0 & 6 & -1 & -6 & -15 \end{array} \right] \quad (2, :) \leftrightarrow (3, :)$$

Now we can complete the job on the second column by subtracting 2 times the second row from the last row.

$$\left[\begin{array}{cccc|c} 1 & 2 & -2 & -7 & -29 \\ 0 & 3 & 0 & -3 & -6 \\ 0 & 0 & 1 & 2 & 11 \\ 0 & 0 & -1 & 0 & -3 \end{array} \right] \quad (4, :) = (4, :) - 2(2, :)$$

Now we shift our attention to the third column. To produce a zero in the entry below the diagonal we must add the third row to the last row.

$$\left[\begin{array}{cccc|c} 1 & 2 & -2 & -7 & -29 \\ 0 & 3 & 0 & -3 & -6 \\ 0 & 0 & 1 & 2 & 11 \\ 0 & 0 & 0 & 2 & 8 \end{array} \right] \quad (4, :) = (4, :) + (3, :)$$

The matrix is now in upper triangular form. Let us find the solution. This last row is shorthand for the equation $2x_4 = 8$. So $x_4 = 2$. The third row now gives $x_3 + 2(4) = 11$, so $x_3 = 3$. The second row gives $3x_2 - 3(4) = -6$ so $x_2 = 2$. Finally the first row gives $x_1 + 2(2) - 2(3) - 7(4) = -29$ so $x_1 = 1$.

Example 3.3 *There is really no need to do anything more in Example 3.2, but let us continue with elementary row operations to put the equations into the "completely solved" form, just to see how this goes.* First we divide the second row by 3.

$$\left[\begin{array}{cccc|c} 1 & 2 & -2 & -7 & -29 \\ 0 & 1 & 0 & -1 & -2 \\ 0 & 0 & 1 & 2 & 11 \\ 0 & 0 & 0 & 2 & 8 \end{array} \right] \quad (2, :) = (1/3)(2, :)$$

Now we subtract twice the second row from the first row.

$$\left[\begin{array}{cccc|c} 1 & 0 & -2 & -5 & -25 \\ 0 & 1 & 0 & -1 & -2 \\ 0 & 0 & 1 & 2 & 11 \\ 0 & 0 & 0 & 2 & 8 \end{array} \right] \quad (1, :) = (1, :) - 2(2, :)$$

Now add twice the third row to the first row. Then divide the last row by 2.

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & -3 \\ 0 & 1 & 0 & -1 & -2 \\ 0 & 0 & 1 & 2 & 11 \\ 0 & 0 & 0 & 1 & 4 \end{array} \right] \quad \begin{array}{l} (1, :) = (1, :) + 2(3, :) \\ (4, :) = (1/2)(4, :) \end{array}$$

Finally, we add various multiples of the last row to the previous rows.

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \end{array} \right] \quad \begin{array}{l} (1, :) = (1, :) + (4, :) \\ (2, :) = (2, :) + (4, :) \\ (3, :) = (3, :) - 2(4, :) \end{array}$$

We now can read the solution off from the last column.

In the previous example there was a unique solution to the system of equations. We already know, from the geometrical meaning of the equations, that sometimes there will be lots of solutions depending on a parameter. This is expected to happen when there are fewer equations than unknowns (e.g., the intersections of two planes in three dimensional space is usually a line) but will also occur in certain degenerate cases when the number of equations is equal to or more than the number of unknowns (e.g., three, or even four, planes may intersect in a line too). What happens in the procedure of row reductions when there are parameters in the solution? Let us look at another example.

Example 3.4 Consider the system

$$\left[\begin{array}{cccc|c} 1 & 3 & 2 & -2 & -1 \\ 1 & 3 & 4 & -2 & 3 \\ -2 & -6 & -4 & 5 & 5 \\ -1 & -3 & 2 & 1 & 6 \end{array} \right]$$

Perform Gaussian Elimination on this system. We begin, as before, by trying to produce zeros in the first column under the diagonal entry. This procedure yields

$$\left[\begin{array}{cccc|c} 1 & 3 & 2 & -2 & -1 \\ 0 & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 4 & -1 & 5 \end{array} \right] \quad \begin{array}{l} (2, :) = (2, :) - (1, :) \\ (3, :) = (3, :) + 2(1, :) \\ (4, :) = (4, :) + (1, :) \end{array}$$

As in the previous example, there is now a zero sitting in the diagonal spot in the second column. Last time, we swapped rows at this point to put a non-zero entry in this place. But now, all the other entries below this one are zero too! So there is nothing we can swap in to save the situation. (Clearly, swapping the first row down is not a good idea, since that would destroy the zero in the first column.) So we just have to admit defeat, and shift our attention one column

to the right. We subtract twice the second row from the fourth row.

$$\left[\begin{array}{cccc|c} 1 & 3 & 2 & -2 & -1 \\ 0 & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 & -3 \end{array} \right] \quad (4, :) = (4, :) - 2(2, :)$$

Now we complete the job by adding the third row to the last row.

$$\left[\begin{array}{cccc|c} 1 & 3 & 2 & -2 & -1 \\ 0 & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \quad (4, :) = (4, :) + (3, :)$$

What are the solutions? The third equation says $x_4 = 3$ and the second equation says $x_3 = 2$. There is nothing new here. However the first equation introduces not just one, but two new variables x_1 and x_2 . It reads $x_1 + 3x_2 + 2(2) - 2(3) = -1$, or, $x_1 + 3x_2 = 1$. Clearly, there are infinitely many values of x_1 and x_2 that satisfy this equation. In fact, if we fix x_2 to be any arbitrary value $x_2 = s$, and then set $x_1 = 1 - 3s$, x_1 and x_2 will be solutions. So for any choice of s

$$x_1 = 1 - 3s, x_2 = s, x_3 = 2, x_4 = 3$$

is a solution. There are infinitely many solutions depending on a parameter s . We could also write this as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix} + s \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

and recognize the solutions as a line in four dimensional space passing through $[1, 0, 2, 3]$ in the direction $[-3, 1, 0, 0]$.

There is one situation left to consider, namely when there are no solutions at all. Geometrically, this happens, for example, when we are trying to find the intersection of two parallel planes. Let us look at an example.

Example 3.5 *Perform Gaussian Elimination on the following system*

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 1 & 4 & 2 \\ -1 & -3 & 0 \\ 2 & 6 & 4 \end{array} \right]$$

We begin in the usual way.

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{array} \right] \quad \begin{array}{l} (2, :) = (2, :) - (1, :) \\ (3, :) = (3, :) + (1, :) \\ (4, :) = (4, :) - 2(1, :) \end{array}$$

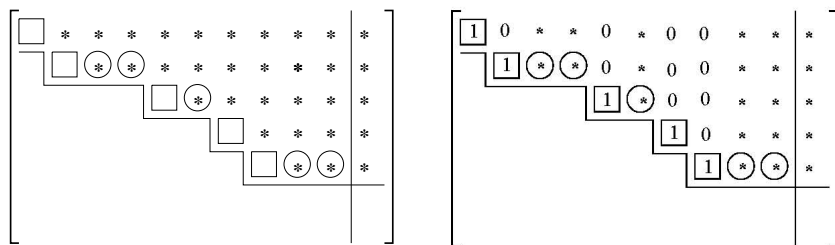


Figure 3.1: After Gaussian Elimination the augmented matrix will be in row echelon form (left). With further work, the augmented matrix can be put in reduced row echelon form (right).

There is nothing left to do in the second column, so we shift our attention to the third column and subtract twice the third row from the fourth row.

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right] \quad (4,:) = (4,:) - 2(3,:)$$

Now we are done. If we write down the equation corresponding to the third row, we get $0x_1 + 0x_2 = 1$, or $0 = 1$. Clearly there is no choice of x_1 or x_2 that makes this true. So this is a system of equations with no solutions.

Let us summarize what we have done in this section. Every system of equations can be brought into upper triangular form using a sequence of elementary row transformations. The resulting upper triangular matrix will look something like what is shown in Figure 3.1 (left). In this diagram, all the entries below the staircase line are zero. The boxes represent non-zero entries. The stars represent arbitrary entries, that may or may not be zero. Each circled star corresponds to a parameter that must be introduced.

If we want to put this example in completely reduced form, we use elementary row operations to zero out the entries lying above the boxes too. Then we multiply each row by a number so that the corner entries (in the boxes) become 1. The completely reduced form for the example above would look like the diagram in Figure 3.1 (right). The official name of this form is the *reduced row echelon form*.

If the bottom of the matrix has a row that is all zeroes, except for the augmented entry, then the system of equations has no solutions. This is because the bottom row stands for an equation of the form $0 = \square$ with $\square \neq 0$. A typical example is shown in Figure 3.2 (left).

If all the steps on the staircase in the non-augmented part of the matrix have size one, then there are no parameters to introduce, and the solution is unique. Notice that in this case there are the same number of equations as variables. A typical example is shown in Figure 3.2 (right).

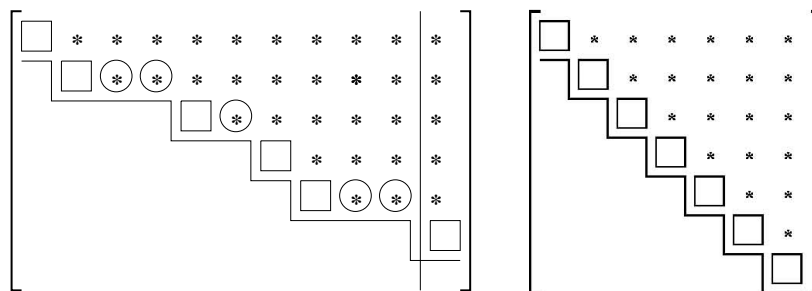


Figure 3.2: Augmented matrices after Gaussian Elimination with no solutions (left) and with a single solution (right).

Finally, we introduce some terminology. The *rank* of a matrix is the number of non-zero rows in the matrix obtained after reducing it to the upper triangular form described above. In other words the rank is the number of boxes in the diagrams above. We can now rephrase the different possibilities in terms of rank. If the rank of the augmented matrix is greater than the rank of the unaugmented matrix (i.e., the matrix without the last column) then there are no solutions. If the rank of the matrix is equal to the number of unknowns then the solution is unique. If the rank r of the matrix is equal to the rank of the unaugmented matrix, but less than the number n of unknowns, then there are $n - r$ parameters in the solution.

Ordinary arithmetic errors are a common source of errors when you do row operations by hand. There is a technique called “the check column” that can often catch such errors. It is described in Section 3.6.1. Another easy way to make sure that we have not made any mechanical errors in calculating the solution of a linear system is just to substitute the purported solution back into the original system and verify that each left hand side really is equal to its corresponding right hand side.

3.2.1 Using MATLAB for row reductions

MATLAB has a built in command called `rref` that reduces a matrix to reduced row echelon form. Let us try it on the example in the previous section. First we define the initial matrix `A`. Remember that the last column of this matrix is the augmented part.

```
A = [1 2 -2 -7 -29; 1 2 -1 -5 -18; 0 3 0 -3 -6; -1 4 1 1 14]
```

To find the reduced row echelon form, simply type

```
>> rref(A)
ans =
1 0 0 0 1
0 1 0 0 2
```

```
0 0 1 0 3
0 0 0 1 4
```

Notice that this MATLAB command did the work of Examples 3.2 and 3.3. The solution to the system can be read off the result of the `rref` command above.

It is important to realize that floating point rounding errors as discussed in Section 2.4.6 can lead to errors in solutions to linear systems computed by MATLAB and other computational tools. At worst, these errors will lead to MATLAB finding “solutions” to problems that do not have exact solutions. In these cases, solutions will often have very large values and often MATLAB will give a warning about the problem being “ill-conditioned”.

3.2.2 Problems

Problem 3.6 Show that the lower triangular system of equations represented by

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 1 & -1 & 0 & 3 \\ 2 & 1 & -8 & -4 \end{array} \right]$$

is also easily solved, by easily solving it! It’s just a matter of convention whether we aim for upper triangular or lower triangular systems in the elimination procedure.

Problem 3.7 The following equations have already been put in upper triangular form. In each case there are infinitely many solutions, depending on one or more parameters. Write down the general expression for the solution in terms of parameters.

$$\left[\begin{array}{cccc|c} 1 & 2 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

$$\left[\begin{array}{cccc|c} 1 & 2 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

$$\left[\begin{array}{cccc|c} 1 & 2 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & 4 \end{array} \right]$$

Problem 3.8 Consider the system of equations represented by the augmented matrix

$$\left[\begin{array}{cccc|c} 1 & 2 & 2 & 2 & 1 \\ 1 & 3 & -1 & 3 & -1 \\ 1 & 0 & 1 & 1 & 5 \\ 0 & 3 & -2 & 2 & -6 \end{array} \right]$$

Show that this set of equations has infinitely many solutions and find a general parametric representation of the solutions.

Problem 3.9 Consider the system of equations represented by the augmented matrix

$$\left[\begin{array}{cccc|c} 1 & 2 & 2 & -7 & 20 \\ 3 & 6 & -3 & -5 & -15 \\ 0 & 6 & 0 & -6 & -10 \\ 2 & -8 & -2 & -2 & 30 \end{array} \right]$$

Put this matrix in upper triangular form and find the solution of the linear system of equations. You can leave your answers in fractions but show the operations you perform in full detail.

Problem 3.10 Solve the following system of equations.

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 2 \\ 2x_1 - 3x_2 + 2x_3 &= 2 \\ 3x_1 + 2x_2 - 4x_3 &= 9 \end{aligned}$$

Problem 3.11 Solve the following system of equations.

$$\begin{aligned} 2x_1 + x_2 - 1x_3 &= 6 \\ x_1 - 2x_2 - 2x_3 &= 1 \\ -x_1 + 12x_2 + 8x_3 &= 7 \end{aligned}$$

Problem 3.12 Solve the following system of equations.

$$\begin{aligned} x_1 + 2x_2 + 4x_3 &= 1 \\ x_1 + x_2 + 3x_3 &= 2 \\ 2x_1 + 5x_2 + 9x_3 &= 1 \end{aligned}$$

Problem 3.13 Solve the following system of equations.

$$\begin{aligned} x_1 + 2x_2 + 4x_3 &= 1 \\ x_1 + x_2 + 3x_3 &= 2 \\ 2x_1 + 5x_2 + 9x_3 &= 3 \end{aligned}$$

Problem 3.14 Solve the following system of equations.

$$\begin{aligned} 3x_1 + x_2 - x_3 + 2x_4 &= 7 \\ 2x_1 - 2x_2 + 5x_3 - 7x_4 &= 1 \\ -4x_1 - 4x_2 + 7x_3 - 11x_4 &= -13 \end{aligned}$$

Problem 3.15 For what values of a , b , c , d , α and β does the system of equations

$$\begin{aligned} ax_1 + bx_2 &= \alpha \\ cx_1 + dx_2 &= \beta \end{aligned}$$

have a unique solution?

Problem 3.16 Consider the system of equations represented by the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 7 \\ 4 & 8 & 6 & 10 \\ -4 & -8 & 10 & 81 \end{array} \right]$$

How many solutions does this linear system of equations have?

Problem 3.17 (Matlab) Consider the following system of equations

$$\begin{aligned} x_1 + 2x_2 + 4x_3 &= 7 \\ 4x_1 + x_2 + 3x_3 &= 2 \\ 0 + 5x_2 + 9x_3 &= a \end{aligned}$$

Write a script in Matlab that generates the augmented matrix and solves the system with the `rref` command for various values of a . Specifically, how does the solution of x_1 vary when you range a from 1 to 10 (equally spaced)?

3.3 Homogeneous Equations

If the coefficients on the right sides of a system of equations are all zero, the system is said to be *homogeneous*. In other words, a homogeneous system is a system of equations of the form

$$\begin{aligned} b_{1,1}x_1 + b_{1,2}x_2 + \cdots + b_{1,n}x_n &= 0 \\ b_{2,1}x_1 + b_{2,2}x_2 + \cdots + b_{2,n}x_n &= 0 \\ \vdots & \\ b_{m,1}x_1 + b_{m,2}x_2 + \cdots + b_{m,n}x_n &= 0 \end{aligned}$$

Given a system of equations, the *associated homogeneous system* is the homogeneous system of equations you get by setting all the right sides to zero.

Geometrically, homogeneous systems describe points, lines and planes that pass through the origin. In fact $\mathbf{x} = \mathbf{0}$, i.e., $x_1 = 0, x_2 = 0, \dots, x_n = 0$ is always a solution to a homogeneous system of equations.

When are there *other (nonzero)* solutions to the above homogeneous system? We have n unknowns and m equations. When we perform the Gaussian reduction, the right-hand sides of the equations will stay zero so the augmented matrix will generally have the form

$$\left[\begin{array}{cccccccc|c} 1 & * & * & * & \cdots & \cdots & \cdots & * & 0 \\ 0 & 1 & * & * & \cdots & \cdots & \cdots & * & 0 \\ 0 & 0 & 0 & 1 & * & \cdots & \cdots & * & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & \cdots & \cdots & \cdots & 1 & * & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

The last several lines may be identically zero. In the last section we saw that there are solutions depending on parameters if the number of variables is greater than the rank of the matrix. Thus, if n (the number of unknowns) is bigger than the number of non-zero lines in the above row-reduced matrix, then there exists a non-zero solution. Otherwise only a trivial solution $x_1 = 0, x_2 = 0, \dots, x_n = 0$ is present. We illustrate the idea with some examples below.

Example 3.6 Consider a homogeneous system

$$\begin{aligned} 3x_1 + 6x_2 + x_3 &= 0 \\ 6x_1 + 2x_2 + 2x_3 &= 0 \\ x_1 + x_2 + 3x_3 &= 0 \end{aligned}$$

The augmented matrix can be reduced by row operations to the form (check!)

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

which implies $x_1 = x_2 = x_3 = 0$. And, in agreement with our above statement, the number of variables (3) is not more than the number of non-zero rows (also 3).

Example 3.7 Consider another homogeneous system:

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

Its augmented matrix is

$$\left[\begin{array}{ccc|c} -1 & 2 & 4 & 0 \\ 2 & -4 & -8 & 0 \\ 1 & -1 & 3 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} -1 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 7 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 10 & 0 \\ 0 & 1 & 7 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right],$$

and the number of nonzero rows is 2, which is less than the number of unknowns, 3. Hence by the above statement there must be a nonzero solution. We find $x_1 = -10x_3$, $x_2 = -7x_3$, with no requirement on x_3 . Hence x_3 is any number t , and we obtain infinitely many nonzero solutions

$$x_1 = -10t, \quad x_2 = -7t, \quad x_3 = t, \quad t \in (-\infty, \infty),$$

one for each value of t .

In a similar manner, if for some homogeneous system with 4 variables the augmented matrix has only 2 nonzero rows, then the general solution has $4-2=2$ free (undefined) variables on which the other two depend.

3.3.1 Properties of solutions of homogeneous systems.

1. A homogeneous system has either one zero-solution ($x_1 = \dots = x_n = 0$) or infinitely-many solutions that depend on parameters.
2. If (x_1, \dots, x_n) and (y_1, \dots, y_n) are solutions to a given homogeneous system, $(x_1 + y_1, \dots, x_n + y_n)$ is also a solution. (Solutions are additive.)
3. If (x_1, \dots, x_n) is a solution to a given homogeneous system, (ax_1, \dots, ax_n) is also a solution, for any number a . (Solutions are scalable.)

The first statement follows from our previous discussion; the other two are easy to verify, using the initial homogeneous system.

3.3.2 Connection of solutions to homogeneous and inhomogeneous systems.

The importance of homogeneous equations comes from the following fact. If $\mathbf{x} = [x_1, x_2, \dots, x_n]$ and $\mathbf{y} = [y_1, y_2, \dots, y_n]$ are two solutions to a (not necessarily homogeneous) system of equations,

$$\begin{array}{ccccccc} b_{1,1}x_1 & + & b_{1,2}x_2 & + \cdots & + & b_{1,n}x_n & = & c_1 \\ b_{2,1}x_1 & + & b_{2,2}x_2 & + \cdots & + & b_{2,n}x_n & = & c_2 \\ \vdots & & \vdots & & & \vdots & & \vdots \\ b_{m,1}x_1 & + & b_{m,2}x_2 & + \cdots & + & b_{m,n}x_n & = & c_m \end{array}$$

then the difference $\mathbf{x} - \mathbf{y} = [x_1 - y_1, x_2 - y_2, \dots, x_n - y_n]$ solves the associated homogeneous system. This is a simple calculation

$$\begin{array}{ccccccc} b_{1,1}(x_1 - y_1) & + & b_{1,2}(x_2 - y_2) & + \cdots & + & b_{1,n}(x_n - y_n) & = & (c_1 - c_1) & = & 0 \\ b_{2,1}(x_1 - y_1) & + & b_{2,2}(x_2 - y_2) & + \cdots & + & b_{2,n}(x_n - y_n) & = & (c_2 - c_2) & = & 0 \\ \vdots & & \vdots & & & \vdots & & \vdots & & \\ b_{m,1}(x_1 - y_1) & + & b_{m,2}(x_2 - y_2) & + \cdots & + & b_{m,n}(x_n - y_n) & = & (c_m - c_m) & = & 0 \end{array}$$

To see the implications of this let us suppose that $\mathbf{x} = \mathbf{q}$ is any particular solution to a (non-homogeneous) system of equations. Then if \mathbf{y} is any other solution $\mathbf{y} - \mathbf{x} = \mathbf{z}$ is a solution of the corresponding homogeneous system. So $\mathbf{y} = \mathbf{q} + \mathbf{z}$. In other words any solution can be written as \mathbf{q} + some solution of the corresponding homogeneous system. Going the other way, if \mathbf{z} is any solution of the corresponding homogeneous system, then $\mathbf{q} + \mathbf{z}$ solves the original system. This can be seen by plugging $\mathbf{q} + \mathbf{z}$ into the equation. So the structure of the set of solutions is

$$\mathbf{x} = \mathbf{q} + (\text{solution to homogeneous system})$$

As you run through all solutions to the homogeneous system on the right, \mathbf{x} runs through all solutions of the original system. Notice that it doesn't matter

which \mathbf{q} you choose as the starting point. This is completely analogous to the parametric form for a line, where the base point can be any point on the line.

If we have applied the process of Gaussian elimination to the original system, and concluded that the general solution has parameters, we will end up with a general solution of the form

$$\mathbf{q} + s_1\mathbf{a}_1 + \cdots + s_n\mathbf{a}_n.$$

Notice that \mathbf{q} is a particular solution (corresponding to all parameters equal to zero) and $s_1\mathbf{a}_1 + \cdots + s_n\mathbf{a}_n$ is the general solution to the corresponding homogeneous system.

These considerations have practical importance if you have to solve a bunch of systems, all with the same coefficients on the left side, but with different coefficients on the right. In this situation, you could first find the general solution to the corresponding homogeneous system of equations. Then to find the general solution to one of the systems, you would only need to find a single particular solution, and then add the general solution to the homogeneous system to obtain all solutions. The only trouble with this is that it might not really be any easier to find a single particular solution than it is to find all solutions.

Example 3.8 Find the general solution of the system of equations

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

In the form $\mathbf{x} = \mathbf{q} + s_1\mathbf{a}_1 + s_2\mathbf{a}_2$. Verify that \mathbf{a}_1 and \mathbf{a}_2 solve the corresponding homogeneous equation. The matrix

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

reduces to

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

so the solutions are $x_1 = 1 - s_1$, $x_2 = s_1$, $x_3 = 2 - 2s_2$, $x_4 = s_2$. This can be written

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix} + s_1 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s_2 \begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

It's easy to check that $\mathbf{a}_1 = [-1, 1, 0, 0]$ and $\mathbf{a}_2 = [0, 0, -2, 1]$ solve the corresponding homogeneous system.

3.3.3 Problems

Problem 3.18 Find the general solution of the system of equations

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 10 \\ -1 & 1 & 1 & 1 & 4 \\ 0 & 1 & 2 & 1 & 14 \end{array} \right]$$

In the form $\mathbf{x} = \mathbf{q} + s_1\mathbf{a}_1 + s_2\mathbf{a}_2$, where \mathbf{a}_1 and \mathbf{a}_2 solve the corresponding homogeneous system of equations.

Problem 3.19 Consider the system of equations

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 4 \\ -1 & -1 & 1 & 2 & -1 \\ 3 & 3 & -1 & -2 & 9 \end{array} \right]$$

Verify that $[4, 0, 3, 0]$ is a solution and write down the general solution.

Problem 3.20 Consider the system of equations given by the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & -3 & 4 & 6 \\ 1 & 9 & -10 & 10 \\ 0 & 6 & -7 & 2 \end{array} \right]$$

Put this matrix in reduced row echelon form and comment on the number of solutions of this system of equations.

3.4 Geometric Applications

Now we will apply Gaussian elimination to some of the geometry problems we studied in the first part of this course.

Let us start with the question of linear independence. Recall that a collection of vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ is called linearly dependent if we can find some non-zero coefficients c_1, c_2, \dots, c_n such that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_n\mathbf{x}_n = \mathbf{0}$$

This is actually a homogeneous system of linear equations for the numbers c_1, \dots, c_n . If $c_1 = c_2 = \dots = c_n = 0$ is the only solution, then the vectors are linearly independent. Otherwise, they are linearly dependent. To decide, we must set up the matrix for the system of equations and perform a row reduction to decide if there is a unique solution or not. In setting up the equations, it is convenient to treat the \mathbf{x}_i 's as column vectors.

Example 3.9 Decide if

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

are linearly independent. The equation $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3 = \mathbf{0}$ can be written

$$\begin{array}{rcccc} c_1 & +c_2 & +c_3 & = & 0 \\ 2c_1 & +c_2 & +2c_3 & = & 0 \\ 0c_1 & +c_2 & +c_3 & = & 0 \end{array}$$

The matrix for this system of equations is

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

Since this is a homogeneous system, we don't have to write the augmented part of the matrix. Performing a row reduction yields

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

Since the number of non-zero rows is the same as the number of variables (three) there are no non-zero solutions. Therefore the vectors are linearly independent.

The row reduction in Example 3.9 also shows that any vector \mathbf{y} in \mathbf{R}^3 can be written as a linear combination of \mathbf{x}_1 , \mathbf{x}_2 and \mathbf{x}_3 . Writing \mathbf{y} as a linear combination of \mathbf{x}_1 , \mathbf{x}_2 and \mathbf{x}_3 means finding coefficients c_1 , c_2 and c_3 such that $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3 = \mathbf{y}$. This is a (non-homogeneous) system of linear equations with augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & y_1 \\ 2 & 1 & 2 & y_2 \\ 0 & 1 & 1 & y_3 \end{array} \right]$$

Using the same Gaussian elimination steps as above, this matrix reduces to

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

where the *'s are some numbers. This system has a (unique) solution.

Example 3.10 *Here is another geometric example. Do the planes whose equations are given by $x_1 + x_2 + x_3 = 1$, $2x_1 + x_2 + 2x_3 = 1$ and $x_2 = 1$ intersect in a single point?* To answer this, we note that the intersection of the three planes is given by the set of points that satisfy all three equations. In other words they satisfy the system of equations whose augmented matrix is

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

A row reduction yields

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

Thus solutions are given by

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

This is the parametric equation of a line. Thus the three planes intersect in a line, not a point.

3.4.1 Problems

Problem 3.21 Are the following vectors linearly dependent or independent?

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 2 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Can every vector in \mathbf{R}^4 be written as a linear combination of these vectors? How about the vector the

$$\mathbf{y}_1 = \begin{bmatrix} 2 \\ 4 \\ -3 \\ 4 \end{bmatrix} ?$$

Problem 3.22 Consider the following vectors \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 such that:

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \mathbf{a}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \mathbf{a}_3 = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$$

Are they linearly independent? Can the vector $\mathbf{y} = \begin{bmatrix} -15 \\ 5 \end{bmatrix}$ be written as a linear combination of \mathbf{a}_1 and \mathbf{a}_2 ?

Problem 3.23 Consider the following 4 dimensional vectors \mathbf{a}_1 , \mathbf{a}_2 , and \mathbf{a}_3 such that

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 0 \\ 4 \\ -3 \end{bmatrix} \quad \mathbf{a}_3 = \begin{bmatrix} 10 \\ 0 \\ -5 \\ 0 \end{bmatrix}$$

Are these linearly independent? Can the vector \mathbf{y} below be written as linear combination of the above three vectors?

Problem 3.24 (Matlab) Consider the following 3 vectors \mathbf{a}_1 , \mathbf{a}_2 , and \mathbf{a}_3 :

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \mathbf{a}_2 = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} \quad \mathbf{a}_3 = \begin{bmatrix} 4 \\ 1 \\ -1 \end{bmatrix}$$

The Matlab command

```
plot3([a,0], [b,0], [c,0])
```

draws a line between $(0,0,0)$ and a point with coordinates (a,b,c) . Write a script that draws the three vectors (use the command `hold on` to be able to overlap them). Are the vectors linearly independent?

3.5 Resistor Networks

3.5.1 Elements of Basic Circuits

Electrical current, often denoted with the variable I , is a measure of charge flow with MKS units of Amperes or Amps (Coulombs per second, where a Coulomb is a measure of electrical charge). The voltage V is an electrical potential measured in Volts (Joules per Coulomb). Thus IV has the units of power (J/s or Watts).

A resistor is a simple electrical component that obeys Ohm's law, that the current through it is proportional to the voltage drop across it. The constant of proportionality is called the resistance R in Ohms (abbreviated Ω , V/A). The current I through a resistor across which there is a voltage drop V satisfies

$$V = IR.$$

The current goes from high to low potential through the resistor.

The resistor networks considered in these notes are circuits with three types of components:

1. Resistors
2. Voltage sources
3. Current sources

The notation we will use for these elements in circuit schematics is shown in Figure 3.3. Later (in Chapter 5) we will see that such a network represents a network with additional elements (inductors and capacitors) at a given instant in time. At a given time, an inductor acts as a current source and a capacitor acts as a voltage source. The current through the capacitor at that instant determines the rate of change of voltage across it, and the voltage across the inductor determines the rate of change of current through it.

In the network, the resistances of all resistors will be given. The voltage across all voltage sources and the currents through all current sources will be given. There are two basic questions to answer about a resistor network with these components.

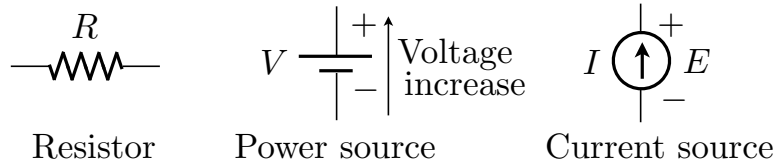


Figure 3.3: Elements in resistor networks.

basic problem: Find the currents through each resistor and each power source and also the voltage drops across each current source. This problem is a *linear system* of equations and so serves as an example of the mathematical techniques learnt in this chapter.

fundamental problem: An important sub-problem is to find the currents through every power source and the voltage drops across each current source. These quantities can be written in terms of the given voltages and current sources directly (eliminating the terms involving resistor currents). Later, solving this problem will tell us how to write a differential equation for the currents through inductors and the voltages across capacitors.

There are two fundamental laws governing the behaviour of circuits that can be used to set up equations that can be solved to answer the questions above. They are Kirchhoff's laws:

1. The sum of voltage drops around any closed loops in the network must be zero.
2. The sum of currents entering a node must be zero.

3.5.2 Two Simple Examples Made Complicated

Consider the simple network with one current source and one resistor shown in Figure 3.4 (left). Clearly, the current through the resistor must be I and the voltage drop across the resistor must be IR using Ohms Law (use the signs in the diagram for the direction of the drop).

Example 3.11 Consider the same example, introducing two nodes into the network as shown in Figure 3.4 (right). We now have three unknowns, V_1 , V_2 and I_2 . Find a linear system for these unknowns and then solve the system. We have made this simple example more complicated but we will learn something as we work through it. Note that by specifying voltages at nodes (which will determine voltage drops across components) we will always satisfy Kirchhoff's first law. Let us write down every other law that applies to this diagram:

$$\begin{aligned}
 I - I_2 &= 0, && \text{Kirchhoff's second law at node 2} \\
 I_2 - I &= 0, && \text{Kirchhoff's second law at node 1} \\
 V_2 - V_1 &= I_2 R, && \text{Ohm's Law over the resistor}
 \end{aligned}$$

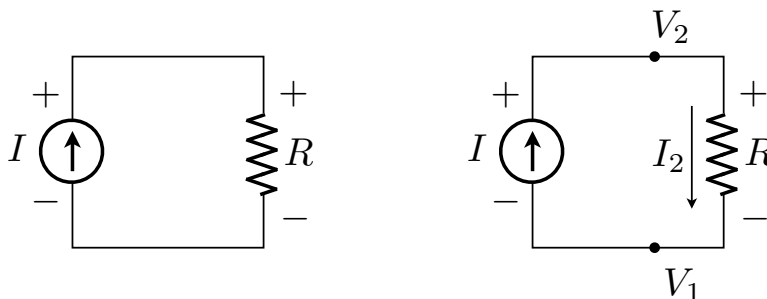


Figure 3.4: The simple resistor network considered in Example 3.11.

If you didn't look too closely, you might be happy thinking these are three equations for the three unknowns I_2 , V_1 and V_2 . However, rewriting gives

$$\begin{aligned} I_2 &= I \\ I_2 &= I \\ V_2 - V_1 - I_2 R &= 0 \end{aligned}$$

In augmented matrix form we can write the system and do Gaussian Elimination:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & I \\ 1 & 0 & 0 & I \\ -R & -1 & 1 & 0 \end{array} \right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & I \\ 0 & -1 & 1 & -RI \\ 0 & 0 & 0 & 0 \end{array} \right]$$

In the augmented matrix above, the unknowns are ordered I_2 , V_1 and then V_2 . The solutions are $I_2 = I$ (expected), $V_2 = s$ and $V_1 = s - RI$ where s is a parameter that can take any value. This seems much more complicated than the intuitive solution at the beginning of this section. However, the conclusions are the same: the current through the resistor is I and the voltage drop across the resistor is

$$V_2 - V_1 = s - (s - RI) = RI.$$

The arbitrary constant in the voltage occurs here because no *reference* voltage has been specified (no point in the circuit has been grounded).

Example 3.12 Consider the circuit in Figure 3.5. Learning from the last example, we have set a reference voltage at the lower left corner of the circuit, and then voltage at the upper left corner is known. The current I_1 is the branch current from the V_2 node to the V_1 node. Form a linear system matching currents and voltages across resistors to Ohm's Law, and matching branch currents at the two nodes. Solve the linear system. There are 5 unknowns V_1 , V_2 , I_1 , I_2 and I_3 in the circuit as shown in the figure (the augmented matrices below will be written with the unknowns in this order). Ohm's law on the four resistors

gives the following linear equations (in order of small to large resistance)

$$\begin{aligned} 12 - V_1 &= I_1 \\ V_1 - V_2 &= 2I_2 \\ V_1 - V_2 &= 3I_3 \\ V_2 &= 4I_1 \end{aligned}$$

and matching the currents at the two nodes gives

$$\begin{aligned} I_1 &= I_2 + I_3 \\ I_2 + I_3 &= I_1 \end{aligned}$$

This gives six equations in five unknowns! (maybe you already see why this happened to us). Writing the six equations above in an augmented matrix gives

$$\left[\begin{array}{ccccc|c} -1 & 0 & -1 & 0 & 0 & -12 \\ 1 & -1 & 0 & -2 & 0 & 0 \\ 1 & -1 & 0 & 0 & -3 & 0 \\ 0 & 1 & -4 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & -1 & 1 & 1 & 0 \end{array} \right] \sim \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 2184/217 \\ 0 & 1 & 0 & 0 & 0 & 1680/217 \\ 0 & 0 & 1 & 0 & 0 & 420/217 \\ 0 & 0 & 0 & 1 & 0 & 252/217 \\ 0 & 0 & 0 & 0 & 1 & 168/217 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

On the left above is the result of Gaussian elimination to reduced row echelon form. The solutions for V_1 , V_2 , I_1 , I_2 and I_3 can be read off the last column of the augmented matrix after reduction ($V_1 = 2184/217$ etc.). Notice that the “extra” equation became the bottom row (all zeros) in the reduced form, which indicates that there was redundant information given in the linear system. If you go back to the expressions for conservation of current at the two nodes above, it is easy to see that these two equations carry the same information.

As the two previous examples show, one has to be a bit careful picking the unknowns and equations in a circuit to get a unique solution without introducing redundant equations. There are several ways to do this and when solving small circuits by hand the linear system can be made much simpler to solve if you pick the “right” technique. In the next section we will describe the “loop current” technique which always leads to a solvable linear system with no extra (redundant) equations.

3.5.3 Loop Currents

We want to be able to see any resistor network and write down equations that will solve it uniquely, with no redundant equations like in the previous example and no non-uniqueness (like that coming from the lack of a reference potential). This can always be done using the following variables: loop currents, that is currents in every elementary loop of the network, and voltage drops across any current sources. This technique is described in more detail below.

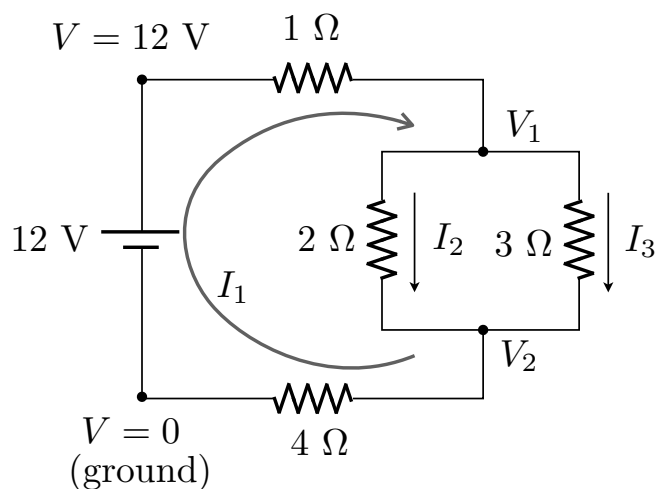


Figure 3.5: The resistor network considered in Example 3.12.

Consider a circuit that can be drawn on a piece of paper with no branches overlapping (a so-called planar network). The branches in the circuit divide the diagram into smaller areas. The set of branches around each of these small areas is called an elementary loop. By assigning a *loop current* to each elementary closed loop of the circuit, the second of Kirchhoff's Circuit Laws is satisfied automatically because in a closed loop the current entering any one point is equal to the current travelling away from that point. Consider Figure 3.7 (left). There are three elementary loops and loop currents i_1 , i_2 and i_3 associated with each of them. Loop currents sum when they overlap in a branch. For example, the current downwards through the $3\ \Omega$ resistor in Figure 3.7 is $i_2 - i_3$. Be careful of signs as you sum loop currents. In the example above, i_2 is downwards through the $3\ \Omega$ resistor but i_3 is upwards, hence it appears with a negative sign in our expression.

In a circuit, it is convenient to take loop currents and voltage drops across current sources as the unknowns. The first step in solving any electric network is to identify the number of elementary loops in the network. If there are m independent loops present, then variables i_1, i_2, \dots, i_m must be introduced to represent the loop currents of each. If there are n current sources, then the variables v_1, v_2, \dots, v_n must be introduced to represent the voltage drop across each source. Together there are $n + m$ unknowns.

We can apply Kirchhoff's voltage law to each of the m loops and obtain m linear equations for the unknowns. The current through each current source must match the loop currents through it. This gives n more linear equations for a total of $n + m$ linear equations for the $n + m$ unknowns.

Example 3.13 *Solve the resistor network in Figure 3.6. Note that this is the same circuit as Example 3.12 but here the loop current method will be used.*

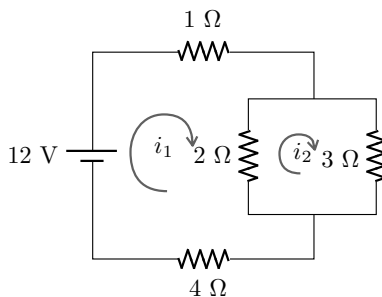


Figure 3.6: The resistor network for Example 3.13

The unknowns are the loop currents i_1 and i_2 (there are no current sources). Remember that the loop currents add in shared components. For example, the current downwards in the 2Ω resistor is $i_1 - i_2$. Using loop currents Kirchhoff's second law is always satisfied. The equations needed to solve for the loop currents are obtained by summing voltage drops around each elementary loop:

$$\begin{aligned} i_1 + 2(i_1 - i_2) + 4i_1 - 12 &= 0, & \text{voltage drops going around loop 1} \\ 3i_2 + 2(i_2 - i_1) &= 0, & \text{voltage drops going around loop 2.} \end{aligned}$$

Collecting terms:

$$\begin{aligned} 7i_1 - 2i_2 &= 12 \\ -2i_1 + 5i_2 &= 0 \end{aligned}$$

which can be solved to give $i_2 = 24/31$ and $i_1 = 420/217$. With these values of i_1 and i_2 the current through each resistor and the power source can be determined, solving the problem.

There are easier ways to solve this *particular* small problem (the easiest is probably to use combinations of the series and parallel resistor laws). However, the loop current rule works for networks of arbitrarily large size and leads to systems of equations with a relatively small number of unknowns. On tests and exams, it is expected that students will be able to apply the idea of loop currents.

Example 3.14 *Solve the network shown in Figure 3.7* There are three independent loop currents which are labelled i_1 , i_2 and i_3 . There is a single current source of $4A$ with voltage drop v across it (minus to plus in the direction of the current). There are also two $10V$ voltage sources. Remember that the current in an electrical branch shared by two loop currents is equal to the (signed) sum of the two loops currents, *i.e.* the current moving to the left through the 5Ω resistor of Figure 3.7 is $i_1 - i_3$ and thus the voltage drop is $5(i_1 - i_3)$. Be careful also of the sign of voltage drops. Moving around loop 1 in the circuit in Figure 3.7 clockwise, when the current source is crossed, there a voltage increase

of v , so this would be a voltage *drop* of $-v$ in the expression for Kirchhoff's second law for this loop written below. We sum the voltage drops around loop 1 beginning at the current source and moving clockwise (the same direction as i_1) to obtain

$$-v + 2i_1 + 5(i_1 - i_3) + 2(i_1 - i_2) = 0$$

which can be simplified to

$$9i_1 - 2i_2 - 5i_3 - v = 0$$

The equations for the voltage drops around loops 2 and 3 are derived similarly

$$\begin{aligned} -2i_1 + 5i_2 - 3i_3 &= -10 \\ -5i_1 - 3i_2 + 8i_3 &= 10 \end{aligned}$$

The final linear equation comes from matching the loop currents to the current source:

$$i_1 = 4$$

The four linear equations above can be solved for the four unknowns i_1, i_2, i_3, v . The solution can be found using MATLAB (the details are in the computer lab #4 guide):

$$\begin{aligned} i_1 &= 4\text{A} \\ i_2 &\approx 2.3871\text{A} \\ i_3 &\approx 4.6452\text{A} \\ v &= 8\text{V} \end{aligned}$$

Now that the loop currents are determined, the branch currents can be written down. For example the current over the 5Ω resistor to the right is $i_3 - i_1 = 0.6452\text{A}$. The right hand panel of Figure 3.7 displays all six branch currents.

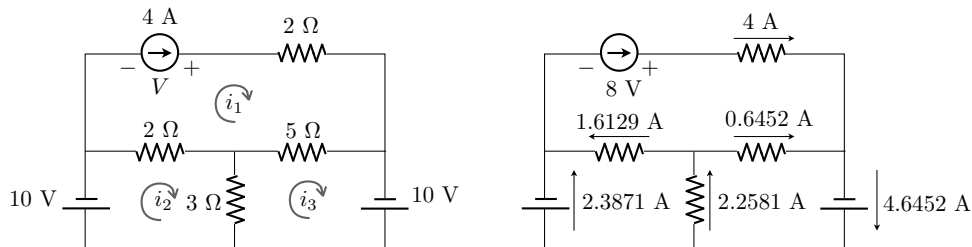


Figure 3.7: The left panel displays the schematic circuit from Example 3.14 with two 10V voltage sources, four resistors and a current source of 4A . The loop currents i_1, i_2, i_3 represent the current in each of the independent closed loops. The right panel is the solution to the electric network on the left panel with branch currents shown.

Example 3.15 *Solve the resistor network in Figure 3.8. In this case, solve both the basic problem when $V = 9$ and $I = 1$, and then the fundamental problem for arbitrary V and I . The unknowns for the problem are i_1 , i_2 , i_3 and E . One equation in the system of unknowns comes from the fact that the loop current variables must match the current source:*

$$i_3 = -I.$$

Note that this equation is so simple we will no longer consider i_3 as a variable but replace i_3 by the known value $-I$ in the equations below. Voltage drops across the three loops give:

$$\begin{aligned} i_1 + 2(i_1 - i_2) + 5i_1 - V &= 0 \\ 3(i_2 - i_3) + 2(i_2 - i_1) &= 0 \\ 5i_3 + E + 3(i_3 - i_2) &= 0 \end{aligned}$$

Since I and V and i_3 (by the discussion above) are known quantities, we move them to the right hand side of the linear equations for i_1 , i_2 and E which are written below

$$\begin{aligned} 8i_1 - 2i_2 &= V \\ -2i_1 + 5i_2 &= -3I \\ -3i_2 + E &= 8I. \end{aligned} \tag{3.1}$$

With $V = 9$ and $I = 1$ this is solved using Gaussian Elimination to give the solution $E = 7\frac{1}{2}$, $i_2 = -1/6$ and $i_1 = 13/12$. Note that the negative value for i_2 means that this loop current physically goes in the opposite direction to that in the Figure. For the fundamental problem, we consider (3.1) for arbitrary values of V and I . We write the system as an augmented matrix and do Gaussian Elimination with symbolic terms in the right hand sides:

$$\left[\begin{array}{ccc|c} 8 & -2 & 0 & V \\ -2 & 5 & 0 & -3I \\ 0 & -3 & 1 & 8I \end{array} \right] \sim \left[\begin{array}{ccc|c} 1 & -1/4 & 0 & V/8 \\ 0 & 1 & 0 & -2/3I + V/18 \\ 0 & 0 & 1 & 6I + V/6 \end{array} \right]$$

so $E = 6I + V/6$ (the voltage across the current source in terms of the given voltage and currents of sources) and (after some algebra) $i_1 = \frac{10}{72}V - \frac{1}{6}I$ (the current through the power source). An alternate approach to solving the fundamental problem is given in next section.

3.5.4 Alternate Presentation of Resistor Networks

Linear systems from resistor networks was presented in a previous version of the notes in a different way. That previous presentation is reproduced here beginning in the next paragraph. This alternate explanation may be helpful to some students.

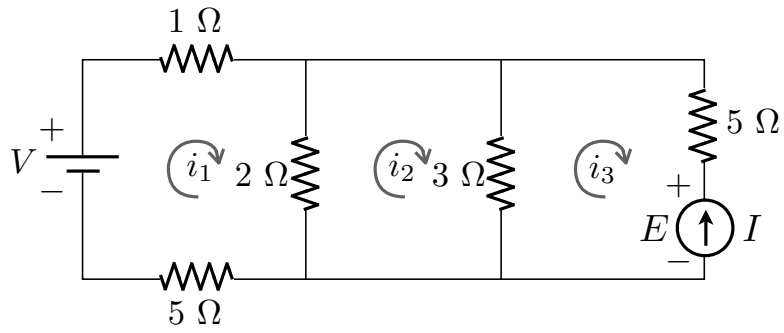


Figure 3.8: The resistor network for Example 3.15

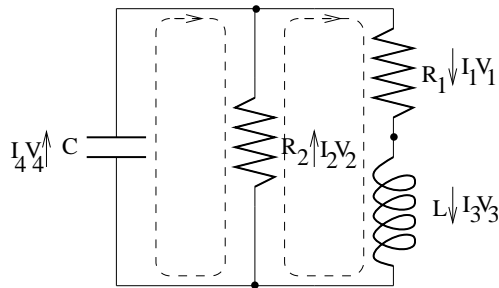


Figure 3.9: A resistor network.

Consider the circuit shown in Figure 3.9. We won't be able to solve this circuit until we have studied differential equations in the last part of this course. However we can make some progress using what we know already.

There are three types of components: resistors, inductors (coils) and capacitors. Associated with each component is the current I flowing through that component, and the voltage drop V across that component. If there are n different components in a circuit, then there are $2n$ variables (currents and voltages) to determine. In the circuit above there are 8.

Of course, these variables are not all independent. They satisfy two types of linear relations: algebraic and differential. We won't touch the differential relations for now, but we can consider the algebraic relations.

The first algebraic relation relates the current and voltage drop across a resistor. If R is the resistance and I and V are the current and voltage drop respectively, then $V = IR$. In our example, this gives two equations

$$\begin{aligned} V_1 &= I_1 R_1 \\ V_2 &= I_2 R_2 \end{aligned}$$

The other two algebraic relations are Kirchhoff's laws. The first of these states that the total voltage drop across any loop in the circuit is zero. For the two loops in the example circuit, this gives the equations

$$\begin{aligned} V_4 - V_2 &= 0 \\ V_1 + V_3 + V_2 &= 0 \end{aligned}$$

Notice we have to take the direction of the arrows into account. The second Kirchhoff law states that current cannot accumulate at a node. At each node, the current flowing in must equal the current flowing out. In the example circuit there are three nodes, giving the equations.

$$\begin{aligned} I_4 + I_2 - I_1 &= 0 \\ I_1 - I_3 &= 0 \\ I_3 - I_2 - I_4 &= 0 \end{aligned}$$

We now want to pick a few of the variables, and solve for all the rest in terms of these. In a small circuit like the example, this can be done "by hand." For example, it's pretty obvious that $I_1 = I_3$ and $V_2 = V_4$ so one could eliminate two variables right off the bat. However, it is also useful to have a systematic way of doing it, that will work for any circuit (but probably will require a computer for anything but the simplest circuit).

As a rule of thumb, you can pick the voltages across the capacitor and the currents across the inductors as basic variables and solve for the rest in terms of these. In other words, we want I_3 and V_4 to be *parameters* when we solve the system of equations. To accomplish this we will choose the order of the

variables with I_3 and V_4 at the end of the list. With this in mind we choose the order $I_1, I_2, I_4, V_1, V_2, V_3, I_3, V_4$. Then the equations become

$$\begin{array}{cccccccc}
 R_1 I_1 & & & & -V_1 & & & = 0 \\
 & R_2 I_2 & & & & -V_2 & & = 0 \\
 & & & & & -V_2 & & +V_4 = 0 \\
 & & & & V_1 & +V_2 & +V_3 & = 0 \\
 -I_1 & +I_2 & +I_4 & & & & & = 0 \\
 I_1 & & & & & & & -I_3 = 0 \\
 & -I_2 & -I_4 & & & & & +I_3 = 0
 \end{array}$$

The matrix for this system is (since it is a homogeneous system of equations, we don't have to bother writing the augmented part)

$$\begin{bmatrix}
 R_1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
 0 & R_2 & 0 & 0 & -1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\
 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
 -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
 0 & -1 & -1 & 0 & 0 & 0 & 1 & 0
 \end{bmatrix}$$

Here is the reduced form of this matrix.

$$\begin{bmatrix}
 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
 0 & 1 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{R_2} \\
 0 & 0 & 1 & 0 & 0 & 0 & -1 & \frac{1}{R_2} \\
 0 & 0 & 0 & 1 & 0 & 0 & -R_1 & 0 \\
 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\
 0 & 0 & 0 & 0 & 0 & 1 & R_1 & 1 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{bmatrix}$$

Thus

$$\begin{aligned}
 I_1 &= I_3 \\
 I_2 &= \frac{1}{R_2} V_4 \\
 I_4 &= I_3 - \frac{1}{R_2} V_4 \\
 V_1 &= R_1 I_3 \\
 V_2 &= V_4 \\
 V_3 &= -R_1 I_3 - V_4
 \end{aligned}$$

So we have succeeded in expressing all the variables in terms of I_3 and V_4 . We therefore need only determine these to solve the circuit completely.

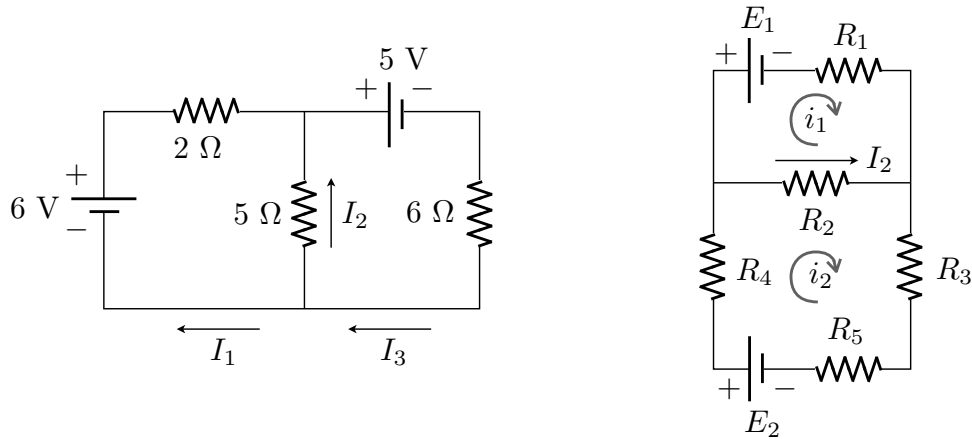


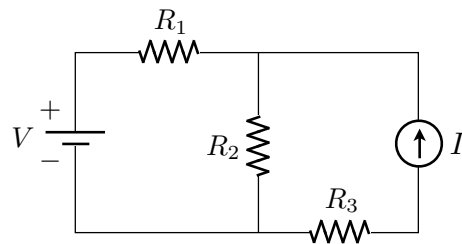
Figure 3.10: Circuit diagrams for Problems 3.25 (left) and 3.26 (right).

3.5.5 Problems

Problem 3.25 Find the currents and voltages in each component of the circuit shown in Figure 3.10 (left).

Problem 3.26 The resistances of the resistors shown in the circuit shown in Figure 3.10 (right) are $R_1 = 4\Omega$, $R_2 = 1\Omega$, $R_3 = R_4 = 2\Omega$. Find the current I_2 that flows through resistor R_2 if the voltage across the batteries are $E_1 = 5V$ and $E_2 = 3V$. In which direction does I_2 flow? Solve the problem as a system of linear equations for the loop currents i_1 and i_2 shown in the diagram.

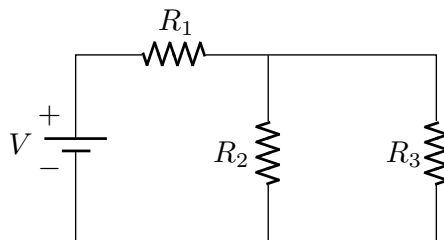
Problem 3.27 Consider the resistor network:



Given $R_1 = 3[\Omega]$, $R_2 = 1[\Omega]$, $R_3 = 4[\Omega]$, $V = 26[V]$ and $I = 2[A]$, answer the following questions:

- What is the voltage drop through R_3 ?
- What is the current flow through R_2 ?
- What is the voltage drop through R_1 ?

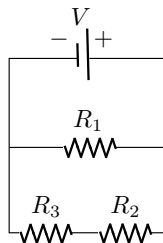
Problem 3.28 Consider the following resistor network:



Suppose that $R_1 = 4[\Omega]$, $R_2 = 1[\Omega]$, $R_3 = 2[\Omega]$ and that the current flow through R_3 is $1.5[A]$. Solve the resistor network and answer the following questions:

- What is the voltage drop across R_3 ?
- What is the current flow through R_2 ?
- What is the voltage drop through V ?

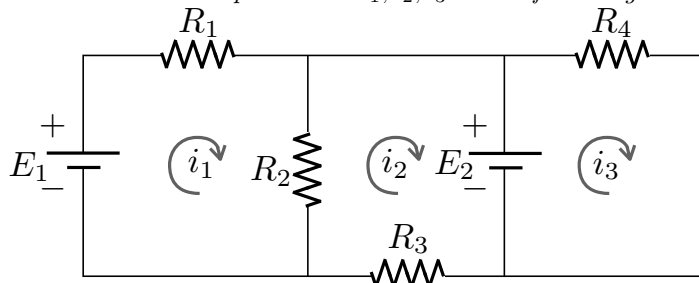
Problem 3.29 Consider the following resistor network:



Suppose that $R_1 = 4[\Omega]$, $R_2 = 2[\Omega]$, $R_3 = 10[\Omega]$ and that $V = 60[V]$. Solve the resistor network and answer the following questions:

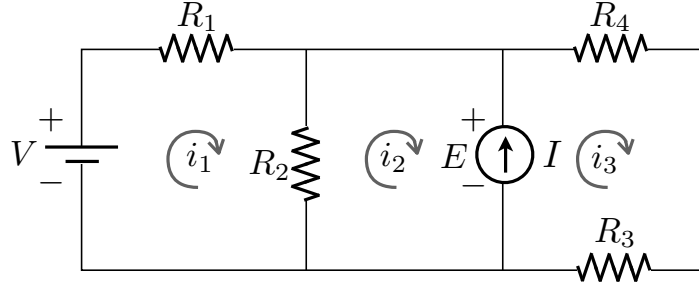
- What is the voltage drop through R_2 ?
- What is the current flow through R_1 ?
- What is the current flow through R_3 ?

Problem 3.30 Find the loop currents i_1, i_2, i_3 in the following network:



where $R_1 = 1[\Omega]$, $R_2 = 3[\Omega]$, $R_3 = 5[\Omega]$, $R_4 = 2[\Omega]$, $E_1 = 10[V]$ and $E_2 = 4[V]$.

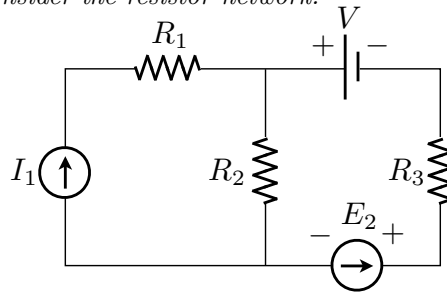
Problem 3.31 Consider the following network:



where $R_1 = 1[\Omega]$, $R_2 = 2[\Omega]$, $R_3 = 1[\Omega]$, $R_4 = 1[\Omega]$, $V = 25[V]$ and $I = 3[A]$.

- Set up and solve the linear system for the loop currents i_1, i_2, i_3 and the voltage E across the current source.
- What is the voltage drop across R_2 ?
- What is the current flow through the voltage source V ?

Problem 3.32 Consider the resistor network:



with $R_1 = 2[\Omega]$, $R_2 = 5[\Omega]$, $R_3 = 3[\Omega]$, $V = 10[V]$, $I_1 = 3[A]$. Suppose that the voltage drop through I_2 is $E_2 = 5[V]$.

- What is the current flow through R_2 ?
- What is the voltage drop through I_1 ?
- What is the current flow through I_1 ?

Problem 3.33 If a circuit contains only resistors, then we can solve it completely using the ideas of this section. Write down the linear equations satisfied by the currents in the circuit shown in Figure 3.11. In this diagram, the component on the far left is a voltage source (battery). The voltage across the voltage source is always E .

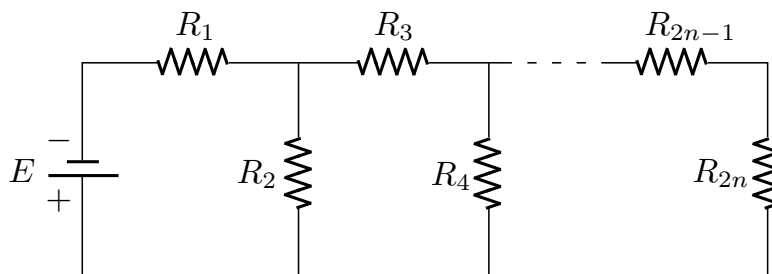


Figure 3.11: The circuit from Problem 3.33.

3.6 Additional Topics

3.6.1 The Check Column

Ordinary arithmetic errors are a big problem when you do row operations by hand. There is a technique called “the check column” (that is modeled after the “parity bit” in computer hardware design) which provides a very effective way to catch mechanical errors. Here is an example which illustrates the technique:

Example 3.16 *The augmented matrix for the system of equations*

$$\begin{aligned} 2x_1 + x_2 + 3x_3 &= 1 \\ 4x_1 + 5x_2 + 7x_3 &= 7 \\ 2x_1 - 5x_2 + 5x_3 &= -7 \end{aligned}$$

is

$$\left[\begin{array}{ccc|c} 2 & 1 & 3 & 1 \\ 4 & 5 & 7 & 7 \\ 2 & -5 & 5 & -7 \end{array} \right]$$

To implement a “check column” you tack onto the right hand side of the augmented matrix an additional column. Each entry in this check column is the sum of all the entries in the row of the augmented matrix that is to the left of the check column entry. For example, the top entry in the check column is $2 + 1 + 3 + 1 = 7$.

$$\left[\begin{array}{ccc|c} 2 & 1 & 3 & 1 & 7 \\ 4 & 5 & 7 & 7 & 23 \\ 2 & -5 & 5 & -7 & -5 \end{array} \right]$$

To use the check column you just perform the same row operations on the check column as you do on the augmented matrix. After each row operation you check that each entry in the check column is still the sum of all the entries in the corresponding row of the augmented matrix.

We now want to eliminate the x_1 's from equations (2) and (3). That is, we want to make the first entries in rows 2 and 3 of the augmented matrix zero. We

can achieve this by subtracting two times row (1) from row (2) and subtracting row (1) from row (3).

$$\begin{array}{l} (1) \\ (2) - 2(1) \\ (3) - (1) \end{array} \left[\begin{array}{ccc|c} 2 & 1 & 3 & 1 \\ 0 & 3 & 1 & 5 \\ 0 & -6 & 2 & -8 \end{array} \right] \begin{array}{l} 7 \\ 9 \\ -12 \end{array}$$

Observe that the check column entry 9 is the sum $0 + 3 + 1 + 5$ of the entries in the second row of the augmented matrix. If this were not the case, it would mean that we made a mechanical error. Similarly the check column entry -12 is the sum $0 - 6 + 2 - 8$.

We have now succeeded in eliminating all of the x_1 's from equations (2) and (3). For example, row 2 now stands for the equation

$$3x_2 + x_3 = 5$$

We next use equation (2) to eliminate all x_2 's from equation (3).

$$\begin{array}{l} (1) \\ (2) \\ (3) + 2(2) \end{array} \left[\begin{array}{ccc|c} 2 & 1 & 3 & 1 \\ 0 & 3 & 1 & 5 \\ 0 & 0 & 4 & 2 \end{array} \right] \begin{array}{l} 7 \\ 9 \\ 6 \end{array}$$

We can now easily solve (3) for x_3 , substitute the result back into (2) and solve for x_2 and so on:

$$\begin{array}{l} (3) \rightarrow 4x_3 = 2 \rightarrow x_3 = \frac{1}{2} \\ (2) \rightarrow 3x_2 + \frac{1}{2} = 5 \rightarrow x_2 = \frac{3}{2} \\ (1) \rightarrow 2x_1 + \frac{3}{2} + 3 \times \frac{1}{2} = 1 \rightarrow x_1 = -1 \end{array}$$

This last step is called “backsolving”.

Note that there is an easy way to make sure that we have not made any mechanical errors in deriving this solution — just substitute the purported solution $(-1, 3/2, 1/2)$ back into the original system:

$$\begin{array}{rclcl} 2(-1) & + & \frac{3}{2} & + & 3 \times \frac{1}{2} & = & 1 \\ 4(-1) & + & 5 \times \frac{3}{2} & + & 7 \times \frac{1}{2} & = & 7 \\ 2(-1) & - & 5 \times \frac{3}{2} & + & 5 \times \frac{1}{2} & = & -7 \end{array}$$

and verify that each left hand side really is equal to its corresponding right hand side. However, if this test fails, all we would know is that we had made an error somewhere in the elimination process. The check column technique can identify the place where the error was made.

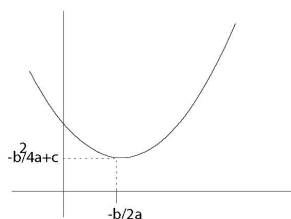


Figure 3.12: The minimization of a quadratic function in one variable.

3.6.2 Quadratic Functions

Let begin by recalling how we would find the minimum of a quadratic function in one variable, namely a parabola given by $f(x) = ax^2 + bx + c$ as shown in Figure 3.12. We simply find the value of x for which the derivative is zero, that is, we solve $f'(x) = 0$. Notice that since f is quadratic, this is a linear equation

$$2ax + b = 0$$

which is easily solved for $x = -b/2a$ (provided $a \neq 0$). So the minimum value is $f(-b/2a) = -b^2/(4a) + c$.

Of course, if a is negative, then the parabola points downwards, and we have found the maximum value, not the minimum value.

A quadratic function of two variables x_1 and x_2 is a function of the form

$$f(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2 + dx_1 + ex_2 + f.$$

(The 2 in front of b is just for convenience.) For what values of x_1 and x_2 is $f(x_1, x_2)$ the smallest? Just like with the parabola in one variable, there may be no such values. It could be that f has a maximum instead, or that f has what is called a *saddle point*. However if f does have a minimum, the procedure described below is guaranteed to find it. (If f has a maximum or saddle point, the procedure will find these points instead.)

The idea behind finding the minimum is simple. Suppose that x_1 and x_2 are the values for which $f(x_1, x_2)$ is smallest. Then the function $g(s) = f(x_1 + s, x_2)$ must have a minimum at $s = 0$. So $g'(0) = 0$. But

$$\begin{aligned} g'(s) &= \frac{d}{ds} f(x_1 + s, x_2) \\ &= \frac{d}{ds} a(x_1 + s)^2 + 2b(x_1 + s)x_2 + cx_2^2 + d(x_1 + s) + ex_2 + f \\ &= 2a(x_1 + s) + 2bx_2 + d \end{aligned}$$

so that the condition is

$$g'(0) = 2ax_1 + 2bx_2 + d = 0.$$

Notice that this expression can be obtained by holding x_2 fixed and differentiating with respect to x_1 . It is called the partial derivative of f with respect to x_1 and is denoted $\frac{\partial f}{\partial x_1}$.

The same argument can be applied to $h(s) = f(x_1, x_2 + s)$ (or $\frac{\partial f}{\partial x_2}$.) This yields

$$h'(0) = \frac{\partial f(x_1, x_2)}{\partial x_2} = 2bx_1 + 2cx_2 + e = 0.$$

Therefore we conclude that the pair of values x_1 and x_2 at which f achieves its minimum satisfy the system of linear equations

$$\begin{aligned} 2ax_1 + 2bx_2 &= -d \\ 2bx_1 + 2cx_2 &= -e \end{aligned}$$

This is a 2 by 2 system with augmented matrix

$$\left[\begin{array}{cc|c} 2a & 2b & -d \\ 2b & 2c & -e \end{array} \right]$$

This is easily generalized to n variables. In this case the quadratic function is given by

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j + \sum_{i=1}^n b_i x_i + c$$

To see this is the same, let us expand out the first term when $n = 2$. Then

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j &= a_{1,1} x_1 x_1 + a_{1,2} x_1 x_2 + a_{2,1} x_2 x_1 + a_{2,2} x_2 x_2 \\ &= a_{1,1} x_1^2 + (a_{1,2} + a_{2,1}) x_1 x_2 + a_{2,2} x_2^2 \end{aligned}$$

So this is just the same as before with $a_{1,1} = a$, $a_{1,2} + a_{2,1} = 2b$ and $a_{2,2} = c$. Notice that we might as well assume that $a_{i,j} = a_{j,i}$, since replacing both $a_{i,j}$ and $a_{j,i}$ with $(a_{i,j} + a_{j,i})/2$ doesn't change f .

If this function f has a minimum we can find it by generalizing the procedure above. In other words we try to find values of x_1, \dots, x_n for which $\partial f / \partial x_1 = \partial f / \partial x_2 = \dots = \partial f / \partial x_n = 0$. This leads to a system of n linear equations whose associated augmented matrix is

$$\left[\begin{array}{cccc|c} 2a_{1,1} & 2a_{1,2} & \dots & 2a_{1,n} & -b_1 \\ 2a_{2,1} & 2a_{2,2} & \dots & 2a_{2,n} & -b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 2a_{n,1} & 2a_{n,2} & \dots & 2a_{n,n} & -b_n \end{array} \right]$$

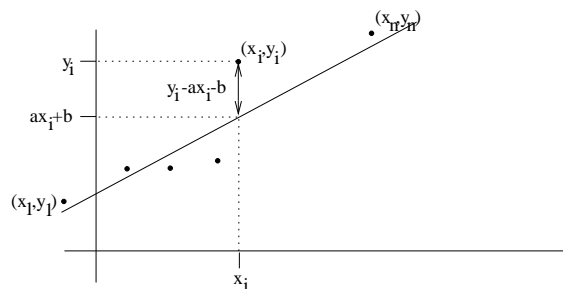


Figure 3.13: Fitting a line through data.

3.6.3 Least squares fit

As a first application let us consider the problem of finding the “best” straight line going through a collection of data points $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. (Careful! the x_i 's are not the unknowns in this problem, but rather the known fixed data points, together with the y_i 's.) Consider Figure 3.13. Which straight line fits best? There is no one answer. One can measure how good the fit of a straight line is in various ways. However the following way of measuring the fit results in a problem that is easy to solve.

Each line is given by an equation $y = ax + b$. So the variables in this problem are a and b . We want to find the values of a and b that give the best fitting line. The vertical distance between the point (x_i, y_i) and the line is given by $|y_i - ax_i - b|$. We will take as a measure of the fit, the square of this quantity, added up over all the data points. So

$$\begin{aligned} f(a, b) &= \sum_i (y_i - ax_i - b)^2 \\ &= \sum_i (y_i^2 + x_i^2 a^2 + b^2 - 2x_i y_i a - 2y_i b + 2x_i ab) \\ &= \left(\sum_i x_i^2 \right) a^2 + 2 \left(\sum_i x_i \right) ab + nb^2 - 2 \left(\sum_i x_i y_i \right) a - 2 \left(\sum_i y_i \right) b + \left(\sum_i y_i^2 \right) \end{aligned}$$

Here we used that $(\sum_i 1) = n$, the number of points. Therefore the linear equations we must solve for a and b are

$$\left[\begin{array}{cc|c} 2 \left(\sum_i x_i^2 \right) & 2 \left(\sum_i x_i \right) & 2 \left(\sum_i x_i y_i \right) \\ 2 \left(\sum_i x_i \right) & 2n & 2 \left(\sum_i y_i \right) \end{array} \right]$$

We could solve these equations numerically in each particular case, but since its just a 2 by 2 system we can also write down the answer explicitly. In fact, the solution to

$$\left[\begin{array}{cc|c} A & B & E \\ C & D & F \end{array} \right]$$

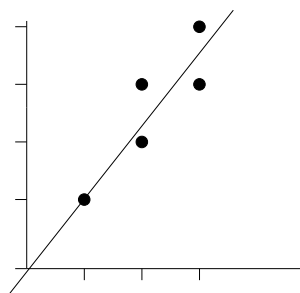


Figure 3.14: The solution in Example 3.17.

is

$$\begin{bmatrix} \frac{DE-BF}{AD-BC} \\ \frac{AF-CE}{AD-BC} \end{bmatrix},$$

provided $AD - BC \neq 0$, as you may check directly, or derive using a sequence of row transformations. So in this case

$$a = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{n(\sum x_i^2) - (\sum x_i)^2}$$

$$b = \frac{(\sum x_i^2)(\sum y_i) - (\sum x_i)(\sum x_i y_i)}{n(\sum x_i^2) - (\sum x_i)^2}$$

Example 3.17 Suppose we want to find the best straight line through the points $(1, 1)$, $(2, 2)$, $(2, 3)$, $(3, 3)$ and $(3, 4)$. Calculate

$$\begin{array}{rcl} \sum 1 & = n & = 5 \\ \sum x_i & = 1 + 2 + 2 + 3 + 3 & = 11 \\ \sum y_i & = 1 + 2 + 3 + 3 + 4 & = 13 \\ \sum x_i^2 & = 1 + 4 + 4 + 9 + 9 & = 27 \\ \sum y_i^2 & = 1 + 4 + 9 + 9 + 16 & = 39 \\ \sum x_i y_i & = 1 + 4 + 6 + 9 + 12 & = 32 \end{array}$$

so

$$a = (5 \cdot 32 - 11 \cdot 13) / (5 \cdot 27 - 11^2) = 17/14 = 1.214 \dots$$

and

$$b = (27 \cdot 13 - 11 \cdot 32) / (5 \cdot 27 - 11^2) = -1/14 = -0.0714 \dots$$

The result is shown in Figure 3.14.

3.6.4 Equilibrium configuration of hanging weights and springs

Consider the problem of n vertically hanging weight connected by springs. What is the equilibrium configuration? We can solve this problem by calculating the

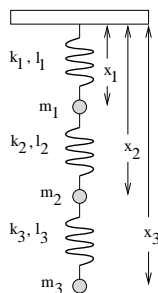


Figure 3.15: Equilibrium configuration of springs.

total potential energy of the system. The equilibrium configuration minimizes the total potential energy. A diagram of the setup is shown in Figure 3.15. Our goal is to compute the numbers x_1, \dots, x_n . In the diagram $n = 3$.

There are two sources of potential energy. One is the potential energy stored in the spring. This is equal to $ks^2/2$, where k is the spring constant that measures the stiffness of the spring, and s is the amount that the spring has been stretched from its natural length. In our problem, suppose that the spring constant of the i th spring is k_i and its natural length is l_i . Then the potential energy stored in the i th spring is $k_i(x_i - x_{i-1} - l_i)^2/2$. To make this formula work out correctly for the first spring we set $x_0 = 0$.

The other source of potential energy is gravity. The gravitational potential energy of the i th weight is $-m_i g x_i$. The reason for the minus sign is that we are measuring distances downward.

Thus the total potential energy in the system for n weights is the function

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \frac{k_i}{2} (x_i - x_{i-1} - l_i)^2 - m_i g x_i.$$

When $n = 3$ this becomes

$$f(x_1, x_2, x_3) = \frac{k_1}{2} (x_1 - l_1)^2 + \frac{k_2}{2} (x_2 - x_1 - l_2)^2 + \frac{k_3}{2} (x_3 - x_2 - l_3)^2 - m_1 g x_1 - m_2 g x_2 - m_3 g x_3$$

This is a quadratic function, so we know how to find the minimum. The equations are obtained by taking partial derivatives: To get the first equation we hold x_2 and x_3 fixed and differentiate with respect to x_1 and so on. Thus the equations are

$$\begin{aligned} k_1(x_1 - l_1) - k_2(x_2 - x_1 - l_2) - m_1 g &= 0 \\ k_2(x_2 - x_1 - l_2) - k_3(x_3 - x_2 - l_3) - m_2 g &= 0 \\ k_3(x_3 - x_2 - l_3) - m_3 g &= 0 \end{aligned}$$

The augmented matrix for this system is

$$\left[\begin{array}{ccc|c} k_1 + k_2 & -k_2 & 0 & m_1 g + k_1 l_1 - k_2 l_2 \\ -k_2 & k_2 + k_3 & -k_3 & m_2 g + k_2 l_2 - k_3 l_3 \\ 0 & -k_3 & k_3 & m_3 g + k_3 l_3 \end{array} \right]$$

Example 3.18 Suppose that the spring constants are $k_1 = 1$, $k_2 = 2$ and $k_3 = 1$. The masses are all equal to 1, $g = 10$ and the natural length of the springs is 1 for all springs (in appropriate units). Find the equilibrium configuration. We must solve

$$\left[\begin{array}{ccc|c} 3 & -2 & 0 & 9 \\ -2 & 3 & -1 & 11 \\ 0 & -1 & 1 & 11 \end{array} \right]$$

Gaussian elimination gives

$$\left[\begin{array}{ccc|c} 3 & -2 & 0 & 9 \\ 0 & -1 & 1 & 11 \\ 0 & 0 & 2 & 106 \end{array} \right]$$

which can be solved to give $x_1 = 31$, $x_2 = 42$, $x_3 = 53$.

3.6.5 Problems

Problem 3.34 Find the “best” straight line going through the points $(1, 1)$, $(2, 1)$, $(2, 3)$, $(3, 4)$, $(3, 5)$ and $(4, 4)$.

Problem 3.35 Consider the problem of finding the parabola $y = ax^2 + bx + c$ that best fits the n data points $(x_1, y_1) \dots (x_n, y_n)$. Derive the system of three linear equations which determine a , b and c . (You need not solve solve them!)

Problem 3.36 Write down the augmented matrix for a system of n weights and springs.

Problem 3.37 Write down the system of equations you would have to solve if there are 5 identical springs with $k_i = 1$ and $l_i = 1$ and five weights with $m_1 = 1$, $m_2 = 2$, $m_3 = 3$, $m_4 = 4$, and $m_5 = 5$.

3.7 Solutions to Chapter Problems

Solution 3.1 The augmented matrix is

$$A = \left[\begin{array}{ccc|c} 1 & -2 & 3 & 6 \\ 4 & -5 & -6 & 7 \\ 8 & 9 & 10 & 11 \end{array} \right]$$

Solution 3.2 Here is the sequence of systems of equations you would get:

$$\begin{array}{rclcl} x_1 & + & x_2 & + & x_3 & = & 6 \\ x_1 & - & x_2 & + & x_3 & = & 0 \\ 2x_1 & + & x_2 & - & 8x_3 & = & -11 \\ \\ x_1 & + & x_2 & + & x_3 & = & 6 \\ & - & 2x_2 & & & = & -6 \\ 2x_1 & + & x_2 & - & 8x_3 & = & -11 \end{array}$$

$$\begin{array}{rclcrcl}
 x_1 & + & x_2 & + & x_3 & = & 6 \\
 & & - & 2x_2 & & = & -6 \\
 & & & - & x_2 & - & 10x_3 & = & -23 \\
 \\
 x_1 & + & x_2 & + & x_3 & = & 6 \\
 & & & & x_2 & = & 3 \\
 & & & - & x_2 & - & 10x_3 & = & -23 \\
 \\
 x_1 & + & x_2 & + & x_3 & = & 6 \\
 & & & & x_2 & = & 3 \\
 & & & & & - & 10x_3 & = & -20
 \end{array}$$

So $x_3 = 2$, $x_2 = 3$ and $x_1 = 1$.

Solution 3.3

$$A = [1, 1, 1, 6; 1, -1, 1, 0; 2, 1, -8, -11]$$

$$A(2, :) = A(2, :) - A(1, :)$$

$$A(3, :) = A(3, :) - 2*A(1, :)$$

$$A(2, :) = -A(2, :)/2$$

$$A(3, :) = A(3, :) + A(2, :)$$

$$A(3, :) = -A(3, :)/10$$

We get the same solution as in the pen-and-paper version in problem 3.2.

Solution 3.4 Perform the sequence of operations to finally end up with the system

$$\begin{array}{rcl}
 2x_3 & = & 4 \\
 3x_1 - 3x_2 + 3x_3 & = & 30 \\
 9x_1 & + & 2x_3 = 26.
 \end{array}$$

Hence, $x_1 = 22/9$, $x_2 = -50/9$, and $x_3 = 2$.

Solution 3.5 We want to solve the system

$$\begin{array}{rcl}
 2x_1 + x_2 & = & 5 \\
 3x_1 + 5x_2 & = & -10.
 \end{array}$$

We shall use the method of substitution.

From the first equation, we solve for x_2 :

$$x_2 = 5 - 2x_1$$

We then substitute it into the second equation:

$$3x_1 + 5(5 - 2x_1) = -10$$

. We now have a decoupled equation, and the solution is $x_1 = 5$, $x_2 = -5$.

Solution 3.6 The first equation reads $x_1 = 3$. The second reads $x_1 - x_2 = 3$, or $3 - x_2 = 3$, or $x_2 = 0$. The third reads $2x_1 + x_2 - 8x_3 = -4$, or $6 + 0 - 8x_3 = -4$ or $x_3 = 5/4$.

Solution 3.7 The last equation gives $x_4 = 2$ and the second last one $x_3 = 2$. Then we have to introduce a parameter $x_2 = s$ and then we find that $x_1 = -5 - 2s$

Thus $\mathbf{x} = \begin{bmatrix} -5 \\ 0 \\ 2 \\ 2 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ In the second system, we have to introduce a

parameter right off the bat. So $x_4 = s_1$ and $x_3 = 4 - s_1$. Moving up one row, we have to introduce another parameter $x_2 = s_2$ and then $x_1 = 1 - 2s_2 - (4 - s_1) -$

$2s_1 = -3 - s_1 - 2s_2$ so $\mathbf{x} = \begin{bmatrix} -3 \\ 0 \\ 4 \\ 0 \end{bmatrix} + s_1 \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} + s_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ The third system is

just the same. The extra rows of zeros have no effect.

Solution 3.8 Perform the following sequence of row operations:

- i) $(2, :) = (2, :) - (1, :)$
- ii) $(3, :) = (3, :) - (1, :)$
- iii) $(3, :) = (3, :) + 2(2, :)$
- iv) $(4, :) = (4, :) - 3(2, :)$
- v) $(4, :) = (4, :) + (3, :)$,

in order to get the reduced row echelon form (RREF)

$$\left[\begin{array}{cccc|c} 1 & 2 & 2 & 2 & 1 \\ 0 & 1 & -3 & 1 & -2 \\ 0 & 0 & -7 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

Thus, we have that the rank of the augmented matrix, which is the same as the rank of the unaugmented matrix is $r = 3$, while the number of unknowns is $n = 4$. Therefore, the system has infinitely many solutions, that can be represented as the general solution

$$\mathbf{x} = \mathbf{q} + s\mathbf{a},$$

where s is any real number, $\mathbf{q} = (q_1, q_2, q_3, q_4)$ is any particular solution of the original system, and $\mathbf{a} = (a_1, a_2, a_3, a_4)$ is any non-zero solution of the corresponding homogeneous system.

From the row echelon form we can see that \mathbf{q} has to satisfy $-7q_3 + q_4 = 0$. Let $q_4 = 7$, then $q_3 = 1$. For the other two q values we have

$$\begin{aligned} q_2 &= -2 + 3q_3 - q_4 = -6 \\ q_1 &= 1 - 2q_2 - 2q_3 - 2q_4 = -3. \end{aligned}$$

Hence, we have that $\mathbf{q} = (-3, -6, 1, 7)$.

Now, to find \mathbf{a} , the homogenous row echelon form is

$$\left[\begin{array}{cccc|c} 1 & 2 & 2 & 2 & 0 \\ 0 & 1 & -3 & 1 & 0 \\ 0 & 0 & -7 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

Thus, \mathbf{a} has to satisfy $-7a_3 + a_4 = 0$. Take $a_4 = 7$, then $a_3 = 1$.

From the second row we have that $a_2 = 3a_3 - a_4 = -4$, and from the first row we get that $a_1 = -2a_2 - 2a_3 - 2a_4 = -8$. So $\mathbf{a} = (-8, -4, 1, 7)$, and a general form of the solution is

$$\mathbf{x} = \begin{bmatrix} -3 \\ -6 \\ 1 \\ 7 \end{bmatrix} + s \begin{bmatrix} -8 \\ -4 \\ 1 \\ 7 \end{bmatrix}$$

Solution 3.9 Perform the following sequence of row operations

i) $(2, :) = (2, :) - 3(1, :)$

ii) $(4, :) = (4, :) - 2(1, :)$

iii) $(4, :) = (4, :) + 2(3, :)$

iv) $(3, :) \leftrightarrow (2, :)$

v) $(4, :) \leftrightarrow (3, :)$

vi) $(4, :) = (4, :) - \frac{3}{2}(3, :)$,

to transform the augmented matrix into the reduced row echelon form

$$\left[\begin{array}{cccc|c} 1 & 2 & 2 & -7 & 20 \\ 0 & 6 & 0 & -6 & -10 \\ 0 & 0 & -6 & 0 & -30 \\ 0 & 0 & 0 & 16 & -30 \end{array} \right].$$

We can now solve the system starting from the bottom:

$$\begin{aligned} x_4 &= -\frac{30}{16} = -\frac{15}{8} \\ x_3 &= \frac{-30}{-6} = 5 \\ x_2 &= \frac{-10 + 6x_4}{6} = -\frac{85}{24} \\ x_1 &= 20 - 2x_2 - 2x_3 + 7x_4 = \frac{95}{24} \end{aligned}$$

Solution 3.10 The matrix

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 2 \\ 2 & -3 & 2 & 2 \\ 3 & 2 & -4 & 9 \end{array} \right]$$

reduces to

$$\left[\begin{array}{ccc|c} 1 & -3 & 3 & 2 \\ 0 & 1 & 4 & -2 \\ 0 & 0 & 19 & 19 \end{array} \right]$$

which has as solution $x_1 = 3$, $x_2 = 2$, $x_3 = 1$.

Solution 3.11 The matrix

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 6 \\ 1 & -2 & -2 & 1 \\ -1 & 12 & 8 & 7 \end{array} \right]$$

reduces to

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 6 \\ 0 & -5/2 & -3/2 & -2 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

which has as solution $x_1 = 13/5 + 4s/5$, $x_2 = 4/5 - 3s/5$, $x_3 = s$.

Solution 3.12 The matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & 1 \\ 1 & 1 & 3 & 2 \\ 2 & 5 & 9 & 1 \end{array} \right]$$

reduces to

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Thus, setting $x_3 = s$, we obtain $x_2 = -1 - s$ and $x_1 = 3 - 2s$.

Solution 3.13 The matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & 1 \\ 1 & 1 & 3 & 2 \\ 2 & 5 & 9 & 3 \end{array} \right]$$

reduces to

$$\left[\begin{array}{ccc|c} 1 & 2 & 4 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 2 \end{array} \right]$$

which has no solutions.

Solution 3.14 The matrix

$$\left[\begin{array}{cccc|c} 3 & 1 & -1 & 2 & 7 \\ 2 & -2 & 5 & -7 & 1 \\ -4 & -4 & 7 & -11 & -13 \end{array} \right]$$

reduces to

$$\left[\begin{array}{cccc|c} 3 & 1 & -1 & 2 & 7 \\ 0 & -8/3 & 17/3 & -25/3 & -11/3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

which has as solution $x_1 = 15/8 - 3s_1/8 + 3s_2/8$, $x_2 = 11/8 + 17s_1/8 - 25s_2/8$, $x_3 = s_1$, $x_4 = s_2$.

Solution 3.15 Probably the easiest way to do this problem is to think geometrically. This system of equations describes the intersection of two lines. The lines will intersect in a single point if they are not parallel. This will happen exactly when the two vectors $[a, b]$ and $[c, d]$ are not parallel. Recall that this can be tested using the determinant. So the equations have a unique solution exactly when $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \neq 0$.

Solution 3.16 Perform the following sequence of row operations:

i) $(2, :) = (2, :) - 4(1, :)$

ii) $(3, :) = (3, :) + 4(1, :)$

iii) $(3, :) = (3, :) + \frac{5}{3}(3, :)$

to get the reduced row echelon matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 7 \\ 0 & 0 & 6 & -18 \\ 0 & 0 & 0 & 139 \end{array} \right].$$

The last equation says $0 = 139$, which cannot happen, therefore this linear system has **zero** solutions.

Solution 3.17 The script is the following:

```
for a = 1:10
A = [1 2 4 7; 4 1 3 2; 0 5 9 a];
rref(A)
end
```

If the value of a changes linearly, then the solution to x_1 (or for that matter, of x_2 and x_3 too) changes linearly with a .

Solution 3.18 Perform the following sequence of row operations:

i) $(2, :) = (2, :) + (1, :)$

$$\text{ii) } (3, :) = (3, :) - (2, :)$$

to get the reduced row echelon matrix

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 10 \\ 0 & 1 & 2 & 1 & 14 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

Thus the row echelon form for the corresponding homogeneous system is

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

Let \mathbf{q} be any solution to the original system. For example, check that $\mathbf{q} = (9, 11, 1, 1)$ solves the system.

Now, let \mathbf{a}_1 and \mathbf{a}_2 be any two linearly independent vectors which solve the corresponding homogeneous system. For example, check that $\mathbf{a}_1 = (1, 1, -1, 1)$, and $\mathbf{a}_2 = (1, -1, -1, 3)$ are two linearly independent vectors, which solve the homogeneous system.

Thus, one representation for the solutions of the system is

$$\mathbf{x} = \begin{bmatrix} 9 \\ 11 \\ 1 \\ 1 \end{bmatrix} + s_1 \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} + s_2 \begin{bmatrix} 1 \\ -1 \\ -1 \\ 3 \end{bmatrix}.$$

Solution 3.19 The general solution is $[4, 0, 3, 0] + s_1[-1, 1, 0, 0] + s_2[0, 0, -2, 1]$.

Solution 3.20 Perform the following sequence of row operations:

$$\text{i) } (2, :) = (2, :) - (1, :)$$

$$\text{ii) } (3, :) = (3, :) - \frac{1}{2}(2, :)$$

$$\text{iii) } (2, :) = \frac{1}{12}(2, :)$$

$$\text{iv) } (1, :) = (1, :) + 3(2, :)$$

to get the reduced row echelon matrix

$$\left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & 7 \\ 0 & 1 & -\frac{7}{6} & 1/3 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

Here the rank of the augmented matrix ($r_a = 2$) is equal to rank of the unaugmented matrix ($r_u = 2$), therefore the linear system has solutions. Since $r_a = 2 < 3$ where 3 is the number of unknowns, there will be an infinite number of solutions with x_3 a free variable.

Solution 3.21 To decide whether the vectors are linearly independent we must decide whether the homogeneous system of equations represented by the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 0 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \text{ has a non zero solution. Row reduction yields } \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}.$$

This shows that the zero solution is unique and therefore the vectors are independent. It can't happen that three vectors span a four dimensional space. To test whether \mathbf{y}_1 is a linear combination of the \mathbf{x}_i 's we try to solve the equation with

$$\text{augmented matrix } \begin{bmatrix} 1 & 1 & 1 & | & 2 \\ 2 & 1 & 0 & | & 4 \\ 0 & -1 & 1 & | & -3 \\ 2 & 1 & 0 & | & 4 \end{bmatrix}. \text{ Row reduction gives } \begin{bmatrix} 1 & 1 & 1 & | & 2 \\ 0 & -1 & -2 & | & 0 \\ 0 & 0 & 3 & | & -3 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}.$$

This system does have a solution, therefore the vector \mathbf{y}_1 is a linear combination of the \mathbf{x}_i 's.

Solution 3.22 Three vectors in the plane cannot be linearly independent.

There are two ways to see that y is a linear combination of a_1 and a_2 :

- a_1 and a_2 are linearly independent because

$$\det \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} = -5$$

Therefore $\{a_1, a_2\}$ form a basis of \mathbb{R}^2 , and every vector is a linear combination of a_1 and a_2 .

- Solve the system

$$x_1 a_1 + x_2 a_2 = y$$

i.e.

$$x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -15 \\ 5 \end{bmatrix}.$$

This system yields the augmented matrix

$$\begin{aligned} \left[\begin{array}{cc|c} 1 & 3 & -15 \\ 2 & 1 & 5 \end{array} \right] &\rightarrow (2, :) = (2, :) - 2(1, :) \rightarrow \left[\begin{array}{cc|c} 1 & 3 & -15 \\ 0 & -5 & 35 \end{array} \right] \\ &\rightarrow (2, :) = -\frac{1}{5}(2, :) \rightarrow \left[\begin{array}{cc|c} 1 & 3 & -15 \\ 0 & 1 & -7 \end{array} \right], \end{aligned}$$

which gives $x_2 = 7$, and $x_1 = 6$.

Solution 3.23 The vectors a_1 , a_2 , and a_3 are linearly independent if the only solution to the system

$$x_1 a_1 + x_2 a_2 + x_3 a_3 = 0$$

is $x_1 = x_2 = x_3 = 0$ So,

$$\begin{aligned} & \left[\begin{array}{ccc|c} 1 & 0 & 10 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 4 & -5 & 0 \\ 0 & -3 & 0 & 0 \end{array} \right] \rightarrow \begin{array}{l} (2,:) = (2,:) - (1,:) \\ (4,:) = -\frac{1}{3}(4,:) \end{array} \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 10 & 0 \\ 0 & 0 & -10 & 0 \\ 0 & 4 & -5 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right] \\ & \rightarrow \begin{array}{l} (2,:) = -\frac{1}{10}(2,:) \\ (3,:) \leftrightarrow (4,:) \end{array} \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 10 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 4 & -5 & 0 \end{array} \right] \rightarrow \begin{array}{l} (2,:) \leftrightarrow (3,:) \\ (4,:) = (4,:) - 4(3,:) + 5(2,:) \end{array} \\ & \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 10 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \end{aligned}$$

The system has a unique solution $x_1 = x_2 = x_3 = 0$, hence the three vectors are linearly independent.

To find if y can be written as a linear combination of the three vectors, we look for a solution of the system

$$\begin{aligned} & \left[\begin{array}{ccc|c} 1 & 0 & 10 & 11 \\ 1 & 0 & 0 & 1 \\ 0 & 4 & -5 & -1 \\ 0 & -3 & 0 & 10 \end{array} \right] \rightarrow \begin{array}{l} (2,:) = (2,:) - (1,:) \\ (3,:) = (3,:) + (4,:) \end{array} \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 10 & 11 \\ 0 & 0 & -10 & -10 \\ 0 & 1 & -5 & 9 \\ 0 & -3 & 0 & 10 \end{array} \right] \rightarrow \\ & \begin{array}{l} (2,:) = -\frac{1}{10}(2,:) \\ (4,:) = (4,:) + 3(3,:) \end{array} \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 10 & 11 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & -5 & 9 \\ 0 & 0 & -15 & 37 \end{array} \right] \end{aligned}$$

The second row of the reduced row echelon matrix implies that $x_3 = 1$, while the third row yields that $x_3 = -\frac{37}{15}$. This contradiction means that the system has no solution, and y is therefore not a linear combination of a_1 , a_2 , and a_3 .

Solution 3.24

```
plot3([1,0],[2,0],[3,0])
hold on
plot3([-1,0],[2,0],[1,0])
plot3([4,0],[1,0],[-1,0])
```

In the figure window, if you go to Tools \rightarrow Rotate 3D, you will see right away that the three vectors do not share the same plane, therefore they must be linearly independent.

Solution 3.25 We apply Kirchhoff's junction rule

$$I_1 + I_2 - I_3 = 0$$

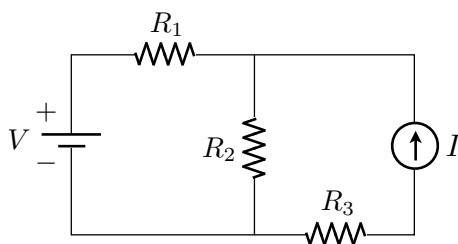


Figure 3.16: Problem 3.27.

and Kirchhoff's loop rules, moving around first the left loop and then the right loop in a clockwise direction:

$$\begin{aligned} -2I_1 + 5I_2 + 6 &= 0 \\ -5 - 6I_3 - 5I_2 &= 0 \end{aligned}$$

Writing these three equations in an augmented matrix with the unknowns ordered I_1 , I_2 and I_3 gives

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ -2 & 5 & 0 & -6 \\ 0 & -5 & -6 & 5 \end{array} \right]$$

which is solved to give $I_1 = 41/52$, $I_2 = -23/26$ and $I_3 = -5/52$ with units of Amperes. The voltage drops are $I_1R_1 = 41/26$, $I_2R_2 = 115/26$ and $I_3R_3 = 15/26$ across the three resistors, with drops in the direction of positive currents. Note that this problem is more easily solved using two loop currents (try it and make sure you get the same solution).

Solution 3.26 We write voltage drops around the two loops in the direction of the loop currents:

$$\begin{aligned} 4i_1 + (i_1 - i_2) + 5 &= 0 \quad \text{or} \quad 5i_1 - i_2 = -5 \\ (i_2 - i_1) + 6i_2 - 3 &= 0 \quad \text{or} \quad -i_1 - i_2 = -5 \end{aligned}$$

which is solved to give $i_2 = 5/17$ and $i_1 = -16/17$. Thus, $I_2 = i_2 - i_1 = 21/17$. Since it is positive, this current flows to the right.

Solution 3.27 We use loop currents to solve the network depicted in figure 3.27.

Kirchhoff's second law gives that $i_2 = -I$, and the first law applied to the closed loop on the left gives

$$R_1i_1 + R_2(i_1 - i_2) = V.$$

As $R_1 = 3[\Omega]$, $R_2 = 1[\Omega]$, $R_3 = 4[\Omega]$, $V = 26[V]$, and $I = 2[A]$, we get

$$i_2 = -2[A] \tag{3.2}$$

$$3i_1 + i_1 + 2 = 26 \tag{3.3}$$

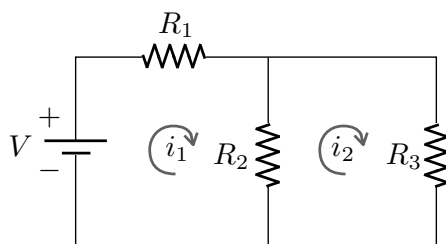


Figure 3.17: Problem 3.28.

Hence $i_1 = 6[A]$, and from the sign we get that i_2 flows in the opposite direction as initially supposed.

- Voltage drop $= R_3 i_2 = -8[V]$. The - sign indicates that the potential is higher at the left of R_3 than on the right.
- The flow through R_2 is $i_1 - i_2 = 8[A]$.
- Voltage drop $= R_1 i_1 = 18[V]$.

Solution 3.28 Again, we use loop currents to solve the problem depicted in figure 3.28.

$$\begin{aligned} R_1 i_1 + R_2(i_1 - i_2) &= V \\ R_3 i_2 + R_2(i_2 - i_1) &= 0, \end{aligned}$$

with $R_1 = 4[\Omega]$, $R_2 = 1[\Omega]$, $R_3 = 2[\Omega]$.

The flow through R_3 is i_2 , which is $1.5[A]$. The system becomes

$$\begin{cases} 4i_1 + (i_1 - 1.5) = V \\ 3 + (1.5 - i_1) = 0 \end{cases}$$

$$\begin{cases} 5i_1 - V = \frac{3}{2} \\ i_1 = \frac{9}{2} \end{cases}$$

Therefore $V = 21[V]$, and $i_1 = 4.5[A]$

- Voltage drop $= R_3 i_2 = 3[V]$.
- Current flow $= i_1 - i_2 = 3[A]$.
- Voltage drop $= V = 21[V]$

Solution 3.29 As before:

$$\begin{cases} R_1(i_1 - i_2) = V \\ R_2 i_2 + R_3 i_2 + R_1(i_2 - i_1) = 0 \end{cases} ,$$

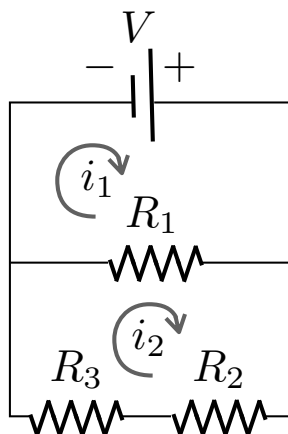


Figure 3.18: Problem 3.29.

with $R_1 = 4[\Omega]$, $R_2 = 2[\Omega]$, $R_3 = 10[\Omega]$, $V = 60[V]$.

Hence, the system becomes

$$\begin{cases} 4(i_1 - i_2) = 15 \\ 2i_2 + 10i_2 + 4(i_2 - i_1) = 0 \end{cases},$$

$$\begin{cases} i_1 - i_2 = 15 \\ 4i_2 - i_1 = 0 \end{cases}.$$

Solving the system we get $i_1 = 20[A]$, $i_2 = 5[A]$.

- Voltage drop $= R_2 i_2 = 10[V]$.
- Current flow through $R_1 = i_1 - i_2 = 15[A]$.
- Current flow through $R_3 = i_2 = 5[A]$.

Solution 3.30 Using Kirchhoff's laws we get the following equations:

$$\begin{cases} R_1 i_1 + R_2(i_1 - i_2) = E_1 \\ R_3 i_2 + R_2(i_2 - i_1) = -E_2 \\ R_4 i_3 = E_2 \end{cases},$$

Plugging in the values of R_1, R_2, R_3, R_4, E_1 , and E_2 , we get

$$\begin{cases} i_1 + 3(i_1 - i_2) = 10 \\ 5i_2 + 3(i_2 - i_1) = -4 \\ 2i_3 = 4 \end{cases},$$

We have that $i_3 = 2$. To find i_1 and i_2 we solve the augmented system

$$\left[\begin{array}{cc|c} 4 & -3 & 10 \\ -3 & 8 & -4 \end{array} \right] \rightarrow (1, :) = (1, :) + (2, :) \rightarrow$$

$$\left[\begin{array}{cc|c} 1 & 5 & 6 \\ -3 & 8 & -4 \end{array} \right] \rightarrow (2, :) = (2, :) + 3(1, :) \rightarrow \left[\begin{array}{cc|c} 1 & 5 & 6 \\ 0 & 23 & 14 \end{array} \right].$$

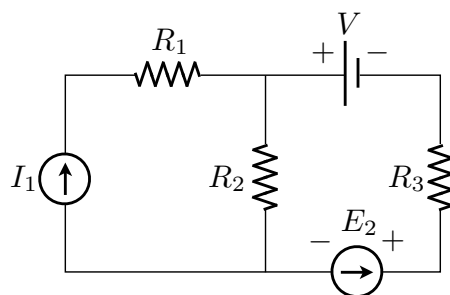


Figure 3.19: Problem 3.32.

So $i_1 = \frac{68}{23}[A]$, $i_2 = \frac{14}{23}[A]$, $i_3 = 2[A]$.

Solution 3.31 From Kirchhoff's first law:

$$I + i_3 = i_2$$

From Kirchhoff's second law:

$$R_1 i_1 + R_2(i_1 - i_2) = V$$

$$E + R_2(i_2 - i_1) = 0$$

$$R_4 i_3 + R_3 i_3 = E$$

Substituting in the actual values, we get

$$\begin{cases} i_1 + 2(i_1 - i_2) = 25 \\ i_3 - i_2 = 3 \\ E + 2(i_2 - i_1) = 0 \\ 2i_3 = E \end{cases}$$

The system, with respect to i_1, i_2, i_3 , and E , yields the following augmented matrix:

$$\left[\begin{array}{cccc|c} 3 & -2 & 0 & 0 & 25 \\ 0 & -1 & 1 & 0 & 3 \\ -2 & 2 & 0 & 1 & 0 \\ 0 & 0 & 2 & -1 & 0 \end{array} \right] \sim \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 11 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 7 \\ 0 & 0 & 0 & 1 & 14 \end{array} \right]$$

a) Then $i_1 = 11[A]$, $i_2 = 4[A]$, $i_3 = 7[A]$, $E = 14[V]$.

b) Voltage drop (from top to bottom): $R_2(i_1 - i_2) = 14[V]$. Can also be seen to be the same as E .

c) Current flow through V : $i_1 = 11[A]$ (upwards).

Solution 3.32 From Kirchhoff's second law:

$$I_1 = i_1$$

$$I_2 = -i_2$$

From Kirchhoff's first law:

$$\begin{aligned} R_1 i_1 + R_2(i_1 - i_2) &= E_1 \\ V + R_3 i_2 + E_2 + R_2(i_2 - i_1) &= 0, \end{aligned}$$

which becomes

$$\begin{cases} i_1 = I_1 \\ i_2 = -3 \\ 7i_1 - 5i_2 = E_1 \\ 10 + 8i_2 - 5i_1 + 5 = 0 \end{cases}$$

We seek i_1 and E_1 :

$$\begin{cases} 7i_1 - E_1 = -15 \\ -5i_1 = 9 \end{cases}.$$

Therefore we have that $i_1 = -\frac{9}{5}[A]$, $E_1 = \frac{12}{5}[V]$.

a) Current flow through R_2 : $i_1 - i_2 = \frac{6}{5}[A]$.

b) Voltage drop through $I_1 = E_1 = \frac{12}{5}[V]$.

c) Current flow through $I_1 = i_1 = -\frac{9}{5}[A]$.

Solution 3.33 Note that this solution follows the alternate description of resistor networks in the notes. Using the identities $V_i = I_i R_i$, the voltage equations for the loops can be written

$$\begin{aligned} I_1 R_1 + I_2 R_2 + E &= 0 \\ I_3 R_3 + I_4 R_4 - I_2 R_2 &= 0 \\ &\vdots \\ I_{2n-1} R_{2n-1} + I_{2n} R_{2n} - I_{2n-2} R_{2n-2} &= 0 \end{aligned}$$

The current equations for the nodes are

$$\begin{aligned} I_1 - I_3 - I_2 &= 0 \\ I_3 - I_5 - I_4 &= 0 \\ &\vdots \\ I_{2n-1} - I_{2n} &= 0 \end{aligned}$$

Solution 3.34 We have $n = 6$, $\sum x_i = 15$, $\sum x_i^2 = 43$, $\sum y_i = 18$, $\sum x_i y_i = 52$, so $a = 14/11 \sim 1.27\dots$ and $b = -2/11 \sim -0.18\dots$

Solution 3.35 We would want to minimize the quadratic function

$$f(a, b, c) = \sum (ax_i^2 + bx_i + c - y_i)^2$$

$$\begin{aligned}
&= \sum (a^2x_i^4 + b^2x_i^2 + c^2 + y_i^2 + 2abx_i^3 + 2acx_i^2 - 2ax_i^2y_i + 2bcx_i - 2bx_iy_i - 2cy_i) \\
&= \left(\sum x_i^4\right)a^2 + \left(\sum x_i^2\right)b^2 + nc^2 + 2\left(\sum x_i^3\right)ab + 2\left(\sum x_i^2\right)ac + 2\left(\sum x_i\right)bc \\
&\quad - 2\left(\sum x_i^2y_i\right)a - 2\left(\sum x_iy_i\right)b - 2\left(\sum y_i\right)c + \left(\sum y_i^2\right)
\end{aligned}$$

The corresponding system of equations is

$$\left[\begin{array}{ccc|c} \left(\sum x_i^4\right) & \left(\sum x_i^3\right) & \left(\sum x_i^2\right) & \left(\sum x_i^2y_i\right) \\ \left(\sum x_i^3\right) & \left(\sum x_i^2\right) & \left(\sum x_i\right) & \left(\sum x_iy_i\right) \\ \left(\sum x_i^2\right) & \left(\sum x_i\right) & n & \left(\sum y_i\right) \end{array} \right]$$

(I've divided each equation by two.)

Solution 3.36 The matrix is

$$\left[\begin{array}{cccccccc|c} k_1 + k_2 & -k_2 & 0 & 0 & \dots & 0 & 0 & 0 & m_1g + k_1l_1 - k_2l_2 \\ -k_2 & k_2 + k_3 & -k_3 & 0 & \dots & 0 & 0 & 0 & m_2g + k_2l_2 - k_3l_3 \\ 0 & -k_3 & k_3 + k_4 & -k_4 & \dots & 0 & 0 & 0 & m_3g + k_3l_3 - k_4l_4 \\ 0 & 0 & -k_4 & k_4 + k_5 & \dots & 0 & 0 & 0 & m_4g + k_4l_4 - k_5l_5 \\ 0 & 0 & 0 & -k_5 & \dots & 0 & 0 & 0 & m_5g + k_5l_5 - k_6l_6 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -k_{n-1} & k_{n-1} + k_n & -k_n & m_{n-1}g + k_{n-1}l_{n-1} - k_nl_n \\ 0 & 0 & 0 & 0 & \dots & 0 & -k_n & k_n & m_g + k_nl_n \end{array} \right]$$

Solution 3.37 The system of equations would be

$$\left[\begin{array}{ccccc|c} 2 & -1 & 0 & 0 & 0 & g \\ -1 & 2 & -1 & 0 & 0 & 2g \\ 0 & -1 & 2 & -1 & 0 & 3g \\ 0 & 0 & -1 & 2 & -1 & 4g \\ 0 & 0 & 0 & -1 & 1 & 5g + 1 \end{array} \right]$$

Chapter 4

Matrices and Determinants

4.1 Matrix operations

A matrix is a rectangular array of numbers. Here is an example of an $m \times n$ matrix.

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

This is sometimes abbreviated $A = [a_{i,j}]$. An $m \times 1$ matrix is called a column vector and a $1 \times n$ matrix is called a row vector. (The convention is that $m \times n$ means m rows and n columns).

Addition and scalar multiplication are defined for matrices exactly as for vectors. If s is a number

$$s \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} = \begin{bmatrix} sa_{1,1} & sa_{1,2} & \cdots & sa_{1,n} \\ sa_{2,1} & sa_{2,2} & \cdots & sa_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ sa_{m,1} & sa_{m,2} & \cdots & sa_{m,n} \end{bmatrix},$$

and

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} \end{bmatrix} = \begin{bmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \cdots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \cdots & a_{2,n} + b_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \cdots & a_{m,n} + b_{m,n} \end{bmatrix}$$

The product of an $m \times n$ matrix $A = [a_{i,j}]$ with a $n \times p$ matrix $B = [b_{i,j}]$ is a $m \times p$ matrix $C = [c_{i,j}]$ whose entries are defined by

$$c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}.$$

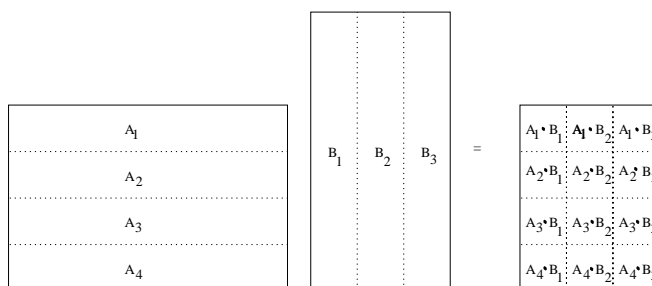


Figure 4.1: Schematic of matrix multiplication as the inner product of rows and columns of the product matrices.

An easy way to remember this is to chop the matrix A into m row vectors of length n and to chop B into p column vectors also of length n , as in the following diagram. The i, j th entry of the product is then the dot product $A_i \cdot B_j$. This is shown schematically in Figure 4.1.

It is *important* to notice that the matrix product AB only makes sense if the number of columns of A equals the number of rows of B . So $A^2 = AA$ only makes sense for a square matrix.

Here is an example

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 4 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 \times 1 + 0 \times 3 + 1 \times 5 + 2 \times 7 & 1 \times 2 + 0 \times 4 + 1 \times 6 + 2 \times 8 \\ 1 \times 1 + 1 \times 3 + 1 \times 5 + 4 \times 7 & 1 \times 2 + 1 \times 4 + 1 \times 6 + 4 \times 8 \\ 0 \times 1 + 0 \times 3 + 1 \times 5 + 1 \times 7 & 0 \times 2 + 0 \times 4 + 1 \times 6 + 1 \times 8 \end{bmatrix} = \begin{bmatrix} 20 & 24 \\ 37 & 44 \\ 12 & 14 \end{bmatrix}$$

Notice that if A is an $m \times n$ matrix, and

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

and

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}$$

Then the equation

$$A\mathbf{x} = \mathbf{b}$$

is a short way of writing the system of linear equations corresponding to the augmented matrix $[A|\mathbf{b}]$.

We will see shortly why matrix multiplication is defined the way it is. For now, you should be aware of some important properties that *don't* hold for matrix multiplication, even though they are true for multiplication of numbers. First of all, in general, AB is *not* equal to BA , even when both products are defined and have the same size. For example, if

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

and

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

then

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

but

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

This example also shows that two non-zero matrices can be multiplied together to give the zero matrix.

Here is a list of properties that do hold for matrix multiplication.

1. $A + B = B + A$
2. $A + (B + C) = (A + B) + C$
3. $s(A + B) = sA + sB$
4. $(s + t)A = sA + tA$
5. $(st)A = s(tA)$
6. $1A = A$
7. $A + \mathbf{0} = A$ (here $\mathbf{0}$ is the matrix with all entries zero)
8. $A - A = A + (-1)A = \mathbf{0}$
9. $A(B + C) = AB + AC$
10. $(A + B)C = AC + BC$
11. $A(BC) = (AB)C$
12. $s(AB) = (sA)B = A(sB)$

4.1.1 MATLAB

Multiplication of matrices can be done using the `*` operator just as for multiplication of scalars. An error results if the matrices are not of compatible size. Powers of matrices can be found using the `^` command like for scalars. The MATLAB command

`A^4`

produces the same result as

`A*A*A*A`

where A is a previously defined, square matrix. Using these commands might be helpful in working out the details of Problem 4.4 below, although you should work out the first few matrix powers by hand for practise. Taking high powers of a matrix will also be helpful in understanding the long time behaviour of random walks described in Section 4.3.

4.1.2 Problems

Problem 4.1 Define

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 1 \end{bmatrix} \quad B = \begin{bmatrix} -1 & 2 \\ -3 & 1 \\ -2 & 1 \end{bmatrix} \quad C = [2 \quad -2 \quad 0] \quad D = \begin{bmatrix} 2 \\ -11 \\ 2 \end{bmatrix}$$

Compute all products of two of these (i.e., AB , AC , etc.) that are defined.

Problem 4.2 Consider the following matrices:

$$\mathbf{A} = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; \quad \mathbf{C} = \begin{bmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{bmatrix}.$$

Compute all the possible products between them.

Problem 4.3 Compute $A^2 = AA$ and $A^3 = AAA$ for

$$A = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}$$

and

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

Problem 4.4 Let

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

- (a) Find A^2 , A^3 and A^4 .
- (b) Find A^k for all positive integers k .
- (c) Find e^{tA} (part of the problem is to invent a reasonable definition!)
- (d) Find a square root of A (i.e., a matrix B with $B^2 = A$).
- (e) Find all square roots of A .

Problem 4.5 Compute A^k for $k = 2, 3, 4$ when

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

4.2 Linear Transformations and Matrices

4.2.1 Linear Transformations

Recall that a function f is a rule that takes an input value x and produces an output value $y = f(x)$. Functions are sometimes called transformations or maps (since they transform, or map, the input value to the output value). In calculus, you have mostly dealt with functions whose input values and output values are real numbers. However, it is also useful to consider functions whose input values and output values are vectors.

We have already encountered this idea when we discussed quadratic functions. A quadratic function such as $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$ can be considered as a transformation (or map) whose input is the vector $\mathbf{x} = [x_1, x_2, x_3]$ and whose output is the number $y = x_1^2 + x_2^2 + x_3^2$. In this case we could write $f(x_1, x_2, x_3)$ as $f(\mathbf{x})$.

An example of a transformation whose inputs and outputs are both vectors in two dimensions is rotation by some angle, say 45° . If \mathbf{x} is the input vector, then the output vector $R(\mathbf{x})$ is the vector you get by rotating \mathbf{x} by 45° in the counter-clockwise direction as shown in Figure 4.2 (left).

A transformation T is called *linear* if for any two input vectors \mathbf{x} and \mathbf{y} and any two numbers s and t ,

$$T(s\mathbf{x} + t\mathbf{y}) = sT(\mathbf{x}) + tT(\mathbf{y}) \quad (4.1)$$

This condition is saying that when we scalar multiply and add two vectors, it doesn't matter whether we (i) do scalar multiplication and addition first and then apply a linear transformation, or (ii) do a linear transformation first and then do scalar multiplication and addition. In both cases we get the same answer. The linearity condition (4.1) is equivalent to the following two conditions:

- (i) For any two vectors \mathbf{x} and \mathbf{y} ,

$$T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}).$$

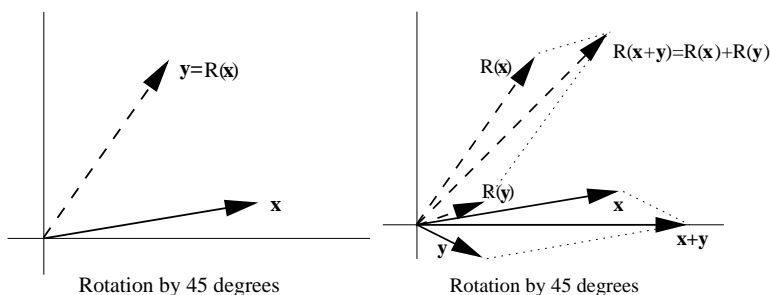


Figure 4.2: Rotation of a vector in 2D (left), graphical evidence that property (i) of linear transformations holds for rotation in 2D (right).

(ii) For any vector \mathbf{x} and any scalar s ,

$$T(s\mathbf{x}) = sT(\mathbf{x})$$

Notice that the quadratic function f above is not a linear transformation, since

$$f(2\mathbf{x}) = (2x_1)^2 + (2x_2)^2 + (2x_3)^2 = 4(x_1^2 + x_2^2 + x_3^2) = 4f(\mathbf{x}).$$

So $f(2\mathbf{x})$ is not equal to $2f(\mathbf{x})$ as would need to be true if f were linear.

However, rotation by 45° is a linear transformation. The picture in Figure 4.2 (right) demonstrates that condition (i) holds.

The most important example of a linear transformation is multiplication by a matrix. If we regard vectors as column vectors, then multiplying an n dimensional vector \mathbf{x} with an $m \times n$ matrix A results in an m dimensional vector $\mathbf{y} = A\mathbf{x}$. The linearity property (4.1) is a consequence of properties 9 and 12 of matrix multiplication listed in Section 4.1. We will see that in fact every linear transformation is of this form.

4.2.2 Rotations in two dimensions

Let us obtain a formula for the transformation that rotates a vector in two dimensions counterclockwise by θ degrees. Let \mathbf{x} be an arbitrary vector. Denote by $\text{Rot}_\theta \mathbf{x}$ the vector obtained by rotating \mathbf{x} counterclockwise by θ degrees. If the angle between \mathbf{x} and the x axis is ϕ , then the components of \mathbf{x} can be written $\mathbf{x} = [x_1, x_2]$ with $x_1 = \|x\| \cos(\phi)$ and $x_2 = \|x\| \sin(\phi)$. This is shown graphically in Figure 4.2.2 (right). To obtain the vector that has been rotated by θ degrees, we simply need to add θ to ϕ in this representation. Thus $\mathbf{y} = \text{Rot}_\theta \mathbf{x} = [y_1, y_2]$, where $y_1 = \|x\| \cos(\phi + \theta)$ and $y_2 = \|x\| \sin(\phi + \theta)$.

To simplify this we can use the addition formulae for sin and cos. Recall that

$$\begin{aligned} \cos(a + b) &= \cos(a) \cos(b) - \sin(a) \sin(b) \\ \sin(a + b) &= \cos(a) \sin(b) + \sin(a) \cos(b) \end{aligned}$$

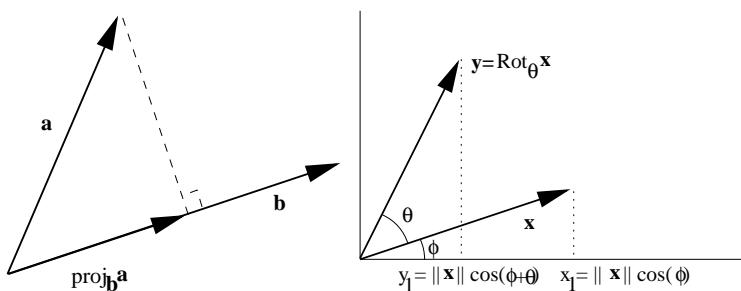


Figure 4.3: Graphical representation of a projection (left). Details of components in a rotation by an angle θ .

Thus

$$\begin{aligned} y_1 &= \|\mathbf{x}\| \cos(\phi + \theta) \\ &= \|\mathbf{x}\| (\cos(\phi) \cos(\theta) - \sin(\phi) \sin(\theta)) \\ &= \cos(\theta)x_1 - \sin(\theta)x_2 \end{aligned}$$

and so

$$\begin{aligned} y_2 &= \|\mathbf{x}\| \sin(\phi + \theta) \\ &= \|\mathbf{x}\| (\sin(\phi) \cos(\theta) + \cos(\phi) \sin(\theta)) \\ &= \sin(\theta)x_1 + \cos(\theta)x_2 \end{aligned}$$

Notice now that this can be written as a matrix product:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

The matrix

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix},$$

also denoted Rot_{θ} , is called a rotation matrix. What this formula is saying is that the linear transformation of rotation by θ degrees is the same as the linear transformation of multiplication by the matrix. In other words, if we want to know the co-ordinates of the vector obtained by rotating \mathbf{x} by θ degrees, we simply calculate $\text{Rot}_{\theta} \mathbf{x}$.

4.2.3 Projections in two dimensions

Now we consider the transformation which projects a vector \mathbf{x} in the direction of another vector \mathbf{a} as shown in Figure 4.2.2 (left). We already have a formula for this transformation. In the *special case* that \mathbf{a} has unit length, the formula is

$$\text{Proj}_{\mathbf{a}} \mathbf{x} = (\mathbf{x} \cdot \mathbf{a}) \mathbf{a}.$$

It follows from the properties of the dot product that

$$\begin{aligned}\text{Proj}_{\mathbf{a}}(s\mathbf{x} + t\mathbf{y}) &= ((s\mathbf{x} + t\mathbf{y}) \cdot \mathbf{a})\mathbf{a} \\ &= (s\mathbf{x} \cdot \mathbf{a} + t\mathbf{y} \cdot \mathbf{a})\mathbf{a} \\ &= s(\mathbf{x} \cdot \mathbf{a})\mathbf{a} + t(\mathbf{y} \cdot \mathbf{a})\mathbf{a} \\ &= s\text{Proj}_{\mathbf{a}}\mathbf{x} + t\text{Proj}_{\mathbf{a}}\mathbf{y}\end{aligned}$$

Thus $\text{Proj}_{\mathbf{a}}$ is a linear transformation. Let us now see that $\text{Proj}_{\mathbf{a}}$ is also given by multiplication by a matrix. If $\mathbf{a} = [a_1, a_2]$ (still considered to be a *unit vector*), then

$$\begin{aligned}\text{Proj}_{\mathbf{a}}\mathbf{x} &= \begin{bmatrix} (x_1a_1 + x_2a_2)a_1 \\ (x_1a_1 + x_2a_2)a_2 \end{bmatrix} \\ &= \begin{bmatrix} a_1^2x_1 + a_1a_2x_2 \\ a_2a_1x_1 + a_2^2x_2 \end{bmatrix} \\ &= \begin{bmatrix} a_1^2 & a_1a_2 \\ a_2a_1 & a_2^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\end{aligned}\tag{4.2}$$

If \mathbf{a} is the unit vector making an angle of θ with the x axis, then $a_1 = \cos(\theta)$ and $a_2 = \sin(\theta)$. Using half angle formulae, we have

$$\begin{aligned}a_1^2 &= \cos^2(\theta) = \frac{1 + \cos(2\theta)}{2} \\ a_2^2 &= \sin^2(\theta) = \frac{1 - \cos(2\theta)}{2} \\ a_1a_2 &= \cos(\theta)\sin(\theta) = \frac{\sin(2\theta)}{2}\end{aligned}$$

Thus the matrix which when multiplied by \mathbf{x} produces the projection of \mathbf{x} onto the line making an angle of θ with the x axis is given by

$$\text{Proj}_{\theta} = \frac{1}{2} \begin{bmatrix} 1 + \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & 1 - \cos(2\theta) \end{bmatrix}$$

Note that if \mathbf{a} is nonzero but *not a unit vector*, then the unit vector

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{\|\mathbf{a}\|} = \frac{(a_1, a_2)}{\sqrt{a_1^2 + a_2^2}}$$

points in the same direction as \mathbf{a} and so (4.2) becomes

$$\text{Proj}_{\mathbf{a}} = \begin{bmatrix} a_1^2/(a_1^2 + a_2^2) & a_1a_2/(a_1^2 + a_2^2) \\ a_2a_1/(a_1^2 + a_2^2) & a_2^2/(a_1^2 + a_2^2) \end{bmatrix}$$

4.2.4 Reflections in two dimensions

A third example of a geometric linear transformation is reflection across a line. The following figure illustrates reflection across a line making an angle θ with the x axis. Let $\text{Ref}_{\theta}\mathbf{x}$ denote the reflected vector as shown in Figure 4.4.

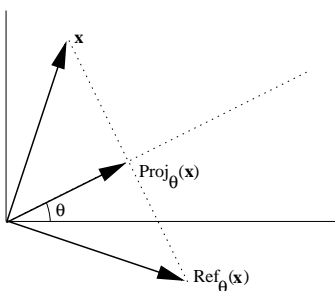


Figure 4.4: Graphical representation of a reflection.

We can obtain the matrix for reflection from the following observation. The vector with tail at \mathbf{x} and head at $\text{Proj}_\theta \mathbf{x}$ is $\text{Proj}_\theta \mathbf{x} - \mathbf{x}$. If we add twice this vector to \mathbf{x} , we arrive at $\text{Ref}_\theta \mathbf{x}$. Therefore

$$\begin{aligned}\text{Ref}_\theta \mathbf{x} &= \mathbf{x} + 2(\text{Proj}_\theta \mathbf{x} - \mathbf{x}) \\ &= 2\text{Proj}_\theta \mathbf{x} - \mathbf{x}\end{aligned}$$

Now if $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, then $I\mathbf{x} = \mathbf{x}$ for any vector \mathbf{x} , since

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1x_1 + 0x_2 \\ 0x_1 + 1x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

I is called the identity matrix.

Now we can write

$$\text{Ref}_\theta \mathbf{x} = 2\text{Proj}_\theta \mathbf{x} - I\mathbf{x} = (2\text{Proj}_\theta - I)\mathbf{x}.$$

This means that the matrix for reflections is $2\text{Proj}_\theta - I$. Explicitly

$$\begin{aligned}\text{Ref}_\theta &= \begin{bmatrix} 1 + \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & 1 - \cos(2\theta) \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}\end{aligned}$$

4.2.5 Every linear transformation is multiplication by a matrix

We have just seen three examples of linear transformations whose action on a vector is given by multiplication by a matrix. Now we will see that for *any* linear transformation $T(\mathbf{x})$ there is a matrix T such that $T(\mathbf{x})$ is the matrix product $T\mathbf{x}$.

To illustrate this suppose that T is a linear transformation that takes three dimensional vectors as input.

Let \mathbf{e}_1 , \mathbf{e}_2 and \mathbf{e}_3 be the standard basis vectors in three dimensions, that is

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Then any vector can be written

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$$

Now, using the linearity property of the linear transformation T , we obtain

$$T(\mathbf{x}) = T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3) = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + x_3T(\mathbf{e}_3)$$

Now take the three vectors $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$ and $T(\mathbf{e}_3)$ and put them in the columns of a matrix which we'll also call T . Then

$$T\mathbf{x} = \left[T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid T(\mathbf{e}_3) \right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = T(\mathbf{e}_1)x_1 + T(\mathbf{e}_2)x_2 + T(\mathbf{e}_3)x_3 = T(\mathbf{x})$$

In other words, the action of the transformation T on a vector \mathbf{x} is the same as

multiplying \mathbf{x} by the matrix $T = \left[T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid T(\mathbf{e}_3) \right]$

The same idea works in any dimension. To find the matrix of a linear transformation $T(\mathbf{x})$ we take the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ (where \mathbf{e}_k has zeros everywhere except for a 1 in the k th spot) and calculate the action of the linear transformation on each one. We then take the transformed vectors $T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_n)$ and put them into the columns of a matrix

$T = \left[T(\mathbf{e}_1) \mid T(\mathbf{e}_2) \mid \dots \mid T(\mathbf{e}_n) \right]$. This matrix T then reproduces the action of the linear transformation, that is, $T(\mathbf{x}) = T\mathbf{x}$.

To see how this works in practise, let's recalculate the matrix for rotations in two dimensions. Under a rotation angle of θ , the vector $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ gets transformed to $T(\mathbf{e}_1) = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$ while the vector $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ gets transformed to $T(\mathbf{e}_2) = \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}$. This is shown graphically in Figure 4.5.

According to our prescription, we must now put these two vectors into the columns of a matrix. This gives the matrix

$$T = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

which is exactly the same as Rot_θ .

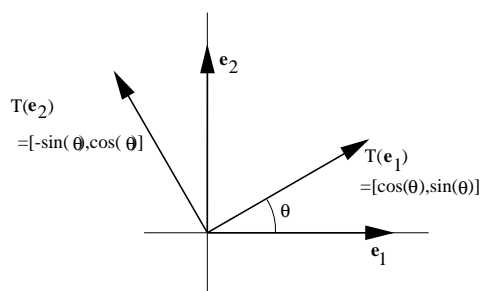


Figure 4.5: Derivation of the matrix representing 2D rotation by rotating coordinate directions.

4.2.6 Composition of linear transformations and matrix product

Suppose we apply one linear transformation T and then follow it by another linear transformation S . For example, think of first rotating a vector and then reflecting it across a line. Then $S(T(\mathbf{x}))$ is again a linear transformation, since

$$S(T(s\mathbf{x} + t\mathbf{y})) = S(sT(\mathbf{x}) + tT(\mathbf{y})) = sS(T(\mathbf{x})) + tS(T(\mathbf{y})).$$

What is the matrix for the composition $S(T(\mathbf{x}))$? We know that there are matrices S and T that reproduce the action of $S(\mathbf{x})$ and $T(\mathbf{x})$. So $T(\mathbf{x})$ is the matrix product $T\mathbf{x}$ and $S(T\mathbf{x})$ is the matrix product $S(T\mathbf{x})$ (here the parenthesis just indicate in which order we are doing the matrix product) But matrix multiplication is associative. So $S(T\mathbf{x}) = (ST)\mathbf{x}$. In other words the matrix for the composition of $S(\mathbf{x})$ and $T(\mathbf{x})$ is simply the matrix product of the corresponding matrices S and T .

For example, to compute the matrix for the transformation of rotation by 45° followed by reflection about the line making an angle of 30° with the x axis we simply compute the product

$$\begin{aligned} \text{Ref}_{30^\circ} \text{Rot}_{45^\circ} &= \begin{bmatrix} \cos(60^\circ) & \sin(60^\circ) \\ \sin(60^\circ) & -\cos(60^\circ) \end{bmatrix} \begin{bmatrix} \cos(45^\circ) & -\sin(45^\circ) \\ \sin(45^\circ) & \cos(45^\circ) \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\sqrt{2}+\sqrt{6}}{4} & \frac{-\sqrt{2}+\sqrt{6}}{4} \\ \frac{\sqrt{6}-\sqrt{2}}{4} & \frac{-\sqrt{6}-\sqrt{2}}{4} \end{bmatrix} \end{aligned}$$

4.2.7 Problems

Problem 4.6 Let \mathbf{a} be a fixed nonzero vector. Show that the transformation $T(\mathbf{x}) = \mathbf{x} + \mathbf{a}$ is not a linear transformation.

Problem 4.7 Let \mathbf{a} be a fixed vector. Show that the transformation $T(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x}$ is a linear transformation (whose output values are numbers).

Problem 4.8 Find the matrices which project on the lines

(a) $x_1 = x_2$

(b) $3x_1 + 4x_2 = 0$

Problem 4.9 Find the matrices that reflect about the lines

(a) $x_1 = x_2$

(b) $3x_1 + 4x_2 = 0$

Problem 4.10 Find the matrices which rotate about the origin in two dimensions by

(a) $\pi/4$

(b) $\pi/2$

(c) π

Problem 4.11 Find the matrix which first reflects about the line making an angle of ϕ with the x axis, and then reflects about the line making an angle of θ with the x axis. Give another geometric interpretation of this matrix.

Problem 4.12 Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear transformation that reflects points across the line $x = 2y$.

1. What is the image of $(1, 10)$, that is, the vector

$$f\left(\begin{bmatrix} 1 \\ 10 \end{bmatrix}\right)?$$

2. Write down the matrix of f .

Problem 4.13 Let $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear transformation that first reflects points across the line $x = -y$ and then rotates points $\pi/2$ radians. Write down the matrix of g .

Problem 4.14 Find the matrix that rotates about the z axis by an angle of θ in three dimensions.

Problem 4.15 Let $T : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ be the map defined by

$$T(x_1, x_2, x_3, x_4) = (x_1 + 4x_2 + 5x_3, 3x_1 - 2x_2 + x_3 - x_4, -x_1 - x_3 + x_4)$$

Show that T is a linear transformation.

Problem 4.16 Let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ given by $T(\mathbf{x}) = A\mathbf{x}$, where

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

Determine whether each given vector is in the range of T . Recall that the range of T is every vector that can “come out” or T , that is every vector in \mathbb{R}^3 that can be written as $T(\mathbf{x})$ for some \mathbf{x} in \mathbb{R}^2 . Note that answering these questions is the same as determining if a linear system has a solution.

1. $\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$.

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

Problem 4.17 Suppose $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ is a linear transformation such that

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad T\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 5 \end{bmatrix}.$$

1. Find $T\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right)$.

2. Write down the matrix of T .

Problem 4.18 Let $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ and $h : \mathbb{R}^2 \rightarrow \mathbb{R}^4$ be linear transformations given by

$$g\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_3 + x_2 \\ x_3 - 2x_2 \end{bmatrix}, \quad h\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ -x_1 \\ x_2 \\ -x_2 \end{bmatrix}.$$

The questions below concern $h \circ g$, the composition of h and g defined by

$$h \circ g(\mathbf{x}) = h(g(\mathbf{x}))$$

for $\mathbf{x} \in \mathbb{R}^3$. Note that this composition is defined because the output of g and the input of h have the same dimension, two.

1. Find $h \circ g\left(\begin{bmatrix} 1 \\ 0 \\ -5 \end{bmatrix}\right)$.

2. Find the matrices of g and h .

3. Find the matrix of $h \circ g$.

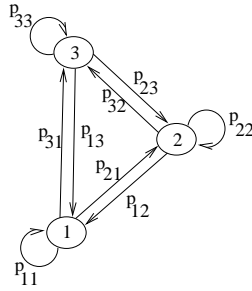


Figure 4.6: Graphical description of a random walk with three states.

4.3 Application: random walks

Consider a system with three states, labelled 1, 2 and 3 in Figure 4.6.

To make the problem more vivid, one can imagine these as being actual locations. A random walker starts off at some location, say location 1 at time 0. Then at a sequence of times, $1, 2, \dots, n, \dots$, the walker either stays where he is, or moves to one of the other locations. The next location is chosen randomly, but according to the transition probabilities $p_{i,j}$. These are numbers between 0 and 1 that measure how likely it is that, starting from location j , the walker will move to location i . If $p_{i,j} = 0$, then there is no chance that the walker will move from j to i , and if $p_{i,j} = 1$, then the walker will move for sure from j to i .

Since the walker must either move from j to another site or stay put, the sum of these probabilities must equal one:

$$p_{1,j} + p_{2,j} + p_{3,j} = \sum_{i=1}^3 p_{i,j} = 1$$

At each time n there is a vector

$$\mathbf{x}_n = \begin{bmatrix} x_{n,1} \\ x_{n,2} \\ x_{n,3} \end{bmatrix}$$

that gives the probabilities that the walker is in location 1, 2 or 3 at time n . Let us compute the vector \mathbf{x}_{n+1} , given \mathbf{x}_n . To start, we must compute $x_{n+1,1}$, the probability that the walker is at location 1 at time $n+1$. There are three ways the walker can end up at location 1. The walker might have been at location 1 at time n and have stayed there. The probability of this is $p_{1,1}x_{n,1}$. Or he might have been at location 2 and have moved to 1. The probability of this is $p_{1,2}x_{n,2}$. Finally, he might have been at location 3 and have moved to 1. The probability of this is $p_{1,3}x_{n,3}$. Thus the total probability that the walker is at location 1 at time $n+1$ is the sum

$$x_{n+1,1} = p_{1,1}x_{n,1} + p_{1,2}x_{n,2} + p_{1,3}x_{n,3}$$

Similarly, for all i

$$x_{n+1,i} = p_{i,1}x_{n,1} + p_{i,2}x_{n,2} + p_{i,3}x_{n,3}$$

But this is exactly the formula for matrix multiplication. So

$$\mathbf{x}_{n+1} = P\mathbf{x}_n$$

where P is the matrix with entries $[p_{ij}]$.

Now suppose the initial probabilities are given by some vector \mathbf{x}_0 . For example, if the walker starts off at location 1, then

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Then after one time step, we have

$$\mathbf{x}_1 = P\mathbf{x}_0$$

Then,

$$\mathbf{x}_2 = P\mathbf{x}_1 = P^2\mathbf{x}_0$$

and so on, so after n time steps

$$\mathbf{x}_n = P^n\mathbf{x}_0$$

Notice that the sum $x_{n,1} + x_{n,2} + x_{n,3} = \sum_{i=1}^3 x_{n,i}$ should equal one, since the total probability of being in one of the three locations must add up to one.

If the initial vector has this property, then it is preserved for all later times. To see this, suppose that $\sum_{j=1}^3 x_j = 1$. Then, since $\sum_{i=1}^3 p_{ij} = 1$

$$\begin{aligned} \sum_{i=1}^3 (Px)_i &= \sum_{i=1}^3 \sum_{j=1}^3 p_{ij}x_j \\ &= \sum_{j=1}^3 \left(\sum_{i=1}^3 p_{ij} \right) x_j \\ &= \sum_{j=1}^3 x_j = 1 \end{aligned}$$

In other words, the vector for the next time step also has components summing to one. Of course, one can generalize this to a system with an arbitrary number of states or locations. Later in the course we will see how to use eigenvalues and eigenvectors to efficiently compute the limit as n tends to infinity of this expression.

A specific but somewhat nerdy example is given below.

Example 4.1 *Ydnew the sorcerer and his apprentice, Xavier, have a magical duel as part of a circus act. They take turns casting spells, which don't always work, but when they do the opponent is knocked out. Xavier always goes first in the duel but his spells only work 1/3 of the time. Ydnew's spells work 1/2 of the time. When they practise, they find that each wins half the time. However, in performances, the duel is limited to three attempts by Xavier. After that, if there has been no knock-out, Ydnew is declared the winner.*

- (a) Describe the duel as a random walk.
- (b) Write the matrix for the random walk.
- (c) Use the matrix to analyze the probability that Ydnew will win the duel when Xavier's attempts are limited and also when they are unlimited.

There are four possible states:

- (1) No winner yet, Xavier's turn
- (2) No winner yet, Ydnew's turn
- (3) Xavier has won.
- (4) Ydnew has won.

These are shown with transition probabilities in Figure 4.7. The transition matrix with the node ordering we have chosen is

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

The duel begins in state (1), that is Xavier's turn, no winner yet:

$$\mathbf{x}^{(0)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The circus duel ends after 3 attempts by Xavier, that is after 5 transitions in the state diagram (five rounds XYXYX). Using MATLAB we compute

$$\mathbf{x}^{(5)} = P^5 \mathbf{x}^{(0)} \approx \begin{bmatrix} 0 \\ 0.0741 \\ 0.4815 \\ 0.4444 \end{bmatrix}.$$

This means that Ydnew will win with probability 0.5185, the sum of the second and fourth components in $\mathbf{x}^{(5)}$ (he can win by a knock-out or if Xavier has

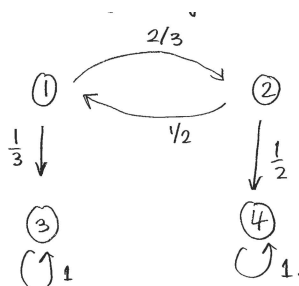


Figure 4.7: Diagram of the sorcerers' duel described in Example 4.1.

not defeated him after his third attempt). To investigate what happens in the unlimited turn version of the duel they use when they practise, you can compute

$$\mathbf{x}^{(n)} = P^n \mathbf{x}^{(0)}$$

for n increasingly large. This gives strong *numerical evidence* that

$$\lim_{n \rightarrow \infty} \mathbf{x}^{(n)} = \begin{bmatrix} 0 \\ 0 \\ 1/2 \\ 1/2 \end{bmatrix}.$$

That is, in the unlimited duel they each win half the time, as they observed. This unlimited case can be analyzed rigorously. The techniques used are not part of the course, but it is done below for completeness. Xavier has a chance to win on any odd turn. On turn 1 he can win if his spell works ($1/3$ chance). On turn 3 Xavier can win if his spell missed on turn 1, Ydnew missed on turn 2 but Xavier succeeds on turn 3 ($\frac{2}{3} \frac{1}{2} \frac{1}{3} = \frac{1}{9}$ chance). On turn 5 he can win if there were failures in turns 1-4 but success on turn 5, with a chance of $\frac{2}{3} \frac{1}{2} \frac{2}{3} \frac{1}{2} \frac{1}{3} = \frac{1}{27}$. Notice the pattern that each successive chance to win goes down by a factor of 3. The total chance of success is the sum of the chances to win at every opportunity:

$$\frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \cdots = \frac{1}{3} \left(1 + \frac{1}{3} + \frac{1}{9} + \cdots \right) = \frac{1}{3} \sum_{i=0}^{\infty} \left(\frac{1}{3} \right)^i = \frac{1}{3} \frac{1}{1 - 1/3} = \frac{1}{2}$$

as predicted from the MATLAB numerical results. To get the second last term above, the expression for the sum of a geometric series was used.

Example 4.2 Investigate a modified sorcerer's duel to the one described in Example 4.1 in which both sorcerers are given a shield which can protect them from one successful spell cast on them. Here, there are 10 possible states

(1) No winner yet, Xavier's turn, both have shields still.

- (2) No winner yet, Xavier's turn, Xavier has lost his shield, but Ydnew still has his.
- (3) No winner yet, Xavier's turn, Xavier still has his shield, but Ydnew has lost his.
- (4) No winner yet, Xavier's turn, both have lost their shields.
- (5-8) Same as (1)-(4) above, but Ydnew's turn.
- (9) Xavier has won.
- (10) Ydnew has won.

Using the ordering of states above, the transition matrix is

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

This matrix can be entered in MATLAB (remember the techniques for entering sparse matrices, that is matrices with mostly zero entries, learnt in your computer lab #2). Starting with $\mathbf{x}^{(0)} = \mathbf{e}_1$ we can compute

$$\mathbf{x}^{(5)} = P^5 \mathbf{x}^{(0)}$$

numerically in MATLAB. Adding components 5-8 (Ydnew wins by default) and component 10 (Ydnew wins by knock-out) gives approximately 0.7778, the chance that Ydnew will win the circus duel. The unlimited duel can be investigated numerically, giving

$$\lim_{n \rightarrow \infty} P^n \mathbf{x}^{(0)} \approx \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0.3750 \\ 0.6250 \end{bmatrix}$$

So in an unlimited match, Ydnew will win about 62.5% of the time.

4.3.1 Problems

Problem 4.19 Consider a random walk with 3 states, where the probability of staying in the same location is zero. Suppose

- the probability of moving from location 1 to location 2 is $1/2$
- the probability of moving from location 2 to location 1 is $1/3$
- the probability of moving from location 3 to location 1 is $1/4$

Write down the matrix P . What is the probability that a walker starting in location 1 is in location 2 after two time steps?

Problem 4.20 Consider a random walk with 3 states, where the probability of staying in the same location is zero. Suppose

- the probability of moving from location 1 to location 2 is $1/3$, that is $p_{2,1} = \frac{1}{3}$.
- the probability of moving from location 2 to location 1 is $2/5$, that is $p_{1,2} = \frac{2}{5}$.
- the probability of moving from location 3 to location 1 is $1/4$, that is $p_{1,3} = \frac{1}{4}$.

1. Write down the matrix P .
2. What is the probability that a walker starting in location 1 is in location 2 after three time steps?
3. What is the probability that a walker starting in location 3 is in location 1 after two time steps?

Problem 4.21 Consider a random walk with 4 states, where all the probabilities $p_{i,j}$ are equal to $1/4$.

1. Compute P and P^n , for every positive integer n .
2. Compute the probability location vectors \mathbf{x}_n ($\mathbf{x}_n = P^n \mathbf{x}_0$) in each of the following cases:

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \text{or} \quad \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Problem 4.22 Suppose that the matrix P given below is the matrix of a random walk:

$$P = \begin{bmatrix} \frac{1}{3} & \frac{1}{4} & \frac{2}{5} \\ 0 & \frac{1}{2} & \frac{1}{5} \\ \frac{2}{3} & \frac{1}{4} & \frac{2}{5} \end{bmatrix}.$$

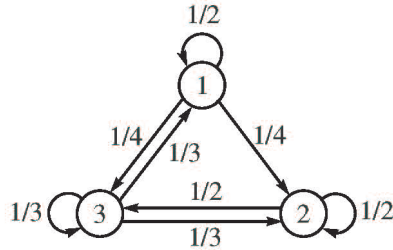


Figure 4.8: Diagram of the random walk in Problem 4.24.

1. What is the probability that a walker starting in location 2 is in location 1 after one time step?
2. What is the probability that a walker starting in location 2 is in location 1 after two time steps?
3. If $\mathbf{x}_0 = \begin{bmatrix} 1/3 \\ 1/3 \\ 2/3 \end{bmatrix}$ what is the probability that the walker is in location 3 after two time steps?

Problem 4.23 Consider a random walk with 3 states, where all the probabilities $p_{i,j}$ are all equal to $1/3$. What is P , P^n ? Compute the probabilities $P^n \mathbf{x}_0$ when

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ (em i.e the walker starts in location 1), } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ (em i.e. the walker}$$

$$\text{starts in location 2), and } \mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ (em i.e. the walker starts in location 3)}$$

Problem 4.24 Consider a random walk with three states with transition probabilities shown in Figure 4.8.

1. Suppose the system starts in state 3. What is the probability that it is in state 2 after 2 steps?
2. Given that

$$P^k \rightarrow \begin{bmatrix} 0.25 & 0.25 & 0.25 \\ 0.375 & 0.375 & 0.375 \\ 0.375 & 0.375 & 0.375 \end{bmatrix}$$

as k tends to infinity, what are the probabilities of the system being in each state after the system has been running for a long time. Show that these probabilities do not depend on the initial state.

4.4 The Transpose

If A is an $m \times n$ matrix, then its transpose A^T is the matrix obtained by flipping A about its diagonal. So the columns of A^T are the rows of A (in the same order) and vice versa. For example, if

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

then

$$A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Another way of saying this is that the i, j th entry of A^T is the same as the j, i th entry of A , that is,

$$a_{i,j}^T = a_{j,i}.$$

There are two important formulae to remember for the transpose of a matrix. The first gives a relation between the transpose and the dot product. If A is an $m \times n$ matrix, then for every $\mathbf{x} \in \mathbf{R}^n$ and $\mathbf{y} \in \mathbf{R}^m$ we have

$$\mathbf{y} \cdot (A\mathbf{x}) = (A^T\mathbf{y}) \cdot \mathbf{x} \quad (4.3)$$

The proof of this formula is a simple calculation.

$$\begin{aligned} \mathbf{y} \cdot (A\mathbf{x}) &= \sum_{i=1}^m y_i \left(\sum_{j=1}^n a_{i,j} x_j \right) \\ &= \sum_{i=1}^m \sum_{j=1}^n y_i a_{i,j} x_j \\ &= \sum_{j=1}^n \sum_{i=1}^m a_{j,i}^T y_i x_j \\ &= \sum_{j=1}^n \left(\sum_{i=1}^m a_{j,i}^T y_i \right) x_j \\ &= (A^T\mathbf{y}) \cdot \mathbf{x} \end{aligned}$$

In fact, the formula (4.3) could be used to define the transpose. Given A , there is exactly one matrix A^T for which (4.3) is true for every \mathbf{x} and \mathbf{y} , and this matrix is the transpose.

The second important formula relates the transpose of a product of matrices to the transposes of each one. For two matrices A and B such that AB is defined the formula reads

$$(AB)^T = B^T A^T \quad (4.4)$$

Notice that the order of the factors is reversed on the right side. To see why (4.4) is true, notice that on the one hand

$$\mathbf{y} \cdot (A\mathbf{B}\mathbf{x}) = ((AB)^T \mathbf{y}) \cdot \mathbf{x}$$

while on the other hand

$$\mathbf{y} \cdot (A\mathbf{B}\mathbf{x}) = \mathbf{y} \cdot (A(\mathbf{B}\mathbf{x})) = (A^T \mathbf{y}) \cdot (\mathbf{B}\mathbf{x}) = (B^T A^T \mathbf{y}) \cdot \mathbf{x}$$

Thus $((AB)^T \mathbf{y}) \cdot \mathbf{x} = (B^T A^T \mathbf{y}) \cdot \mathbf{x}$ for every \mathbf{x} and \mathbf{y} . This can only be true if (4.4) holds.

4.4.1 MATLAB

The MATLAB operator ' can be used to take the transpose of a matrix.

4.4.2 Problems

Problem 4.25 Verify formula (4.3) for $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ and

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

Problem 4.26 What is $(A^T)^T$?

Problem 4.27 Verify (4.4) for

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

Problem 4.28 Show that if A and B are both $m \times n$ matrices such that $\mathbf{y} \cdot (A\mathbf{x}) = \mathbf{y} \cdot (B\mathbf{x})$ for every $\mathbf{y} \in \mathbf{R}^m$ and every $\mathbf{x} \in \mathbf{R}^n$, then $A = B$.

Problem 4.29 Show that if you think of (column) vectors in \mathbf{R}^n as $n \times 1$ matrices then

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y}$$

Now use this formula and (4.4) to derive (4.3).

4.5 Matrix Inverses

To solve the (scalar) equation

$$ax = b$$

for x we simply multiply both sides by $a^{-1} = \frac{1}{a}$. Then, since $a^{-1}a = 1$, we find

$$x = a^{-1}b.$$

Of course, if $a = 0$ this doesn't work, since we cannot divide by zero. In fact, if $a = 0$ the equation $ax = b$ either has no solutions, if $b \neq 0$, or infinitely many solutions (every value of x), if $b = 0$.

We have seen that a system of linear equations can be rewritten

$$A\mathbf{x} = \mathbf{b}$$

where A is a known matrix \mathbf{x} is the unknown vector to be solved for, and \mathbf{b} is a known vector. Suppose we could find an inverse matrix B (analogous to a^{-1}) with the property that $BA = I$ (recall that I denotes the identity matrix). Then we could matrix multiply both sides of the equation by B yielding

$$BA\mathbf{x} = B\mathbf{b}$$

But $BA\mathbf{x} = I\mathbf{x} = \mathbf{x}$, so $\mathbf{x} = B\mathbf{b}$. Thus there is a unique solution and we have a formula for it.

Just as in the numerical case, where a could be zero, we can't expect to find an inverse matrix in all cases. After all, we know that there are linear systems of equations with no solutions and with infinitely many solutions. In these situations there can be no inverse matrix.

When considering matrix inverses, we will always assume that we are dealing with square (i.e., $n \times n$) matrices.

Definition 4.1 *If A is an $n \times n$ matrix, then B is called the inverse of A , and denoted $B = A^{-1}$, if*

$$BA = I$$

where I is the $n \times n$ identity matrix (with each diagonal entry equal to 1 and all other entries 0).

Here is an example. Suppose

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$$

then the inverse matrix is

$$B = A^{-1} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$$

since

$$\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} 6-5 & 3-3 \\ 10-10 & -5+6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This means that to solve the linear equation

$$\begin{aligned} 2x_1 + x_2 &= 2 \\ 5x_1 + 3x_2 &= 4 \end{aligned}$$

we write it as a matrix equation

$$\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

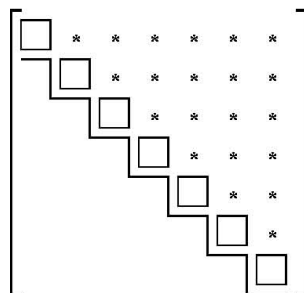


Figure 4.9: Diagram of the reduced form of an invertible matrix.

and then multiply both sides by the inverse to obtain

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

Here is an example of a matrix that doesn't have an inverse. Let

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

To see that A doesn't have an inverse, notice that the homogeneous equations $A\mathbf{x} = \mathbf{0}$ has a non-zero solution $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. If A had an inverse B , then we could multiply both sides of the equation $A\mathbf{x} = \mathbf{0}$ by B to obtain $\mathbf{x} = B\mathbf{0} = \mathbf{0}$. But this is false. Therefore there cannot be an inverse for A .

Clearly, having an inverse is somehow connected to whether or not there are any non-zero solutions of the homogeneous equation $A\mathbf{x} = \mathbf{0}$. Recall that $A\mathbf{x} = \mathbf{0}$ has only the zero solution precisely when $A\mathbf{x} = \mathbf{b}$ has a unique solution for any \mathbf{b} .

Let A be an $n \times n$ matrix. The following conditions are equivalent:

- (1) A is invertible.
- (2) The equation $A\mathbf{x} = \mathbf{b}$ always has a unique solution.
- (3) The equation $A\mathbf{x} = \mathbf{0}$ has as the only solution $\mathbf{x} = \mathbf{0}$.
- (4) The rank of A is n .
- (5) The reduced form of A is as shown in Figure 4.9.

We already know that the conditions (2), (3), (4) and (5) are all equivalent. We also just have seen that if A is invertible with inverse B , then the solution of $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = B\mathbf{b}$ so it exists and since we have a formula for it, it is unique.

So we just have to show that if $A\mathbf{x} = \mathbf{b}$ always has a unique solution, then A has an inverse. Consider the transformation that takes a vector \mathbf{b} to the

unique solution \mathbf{x} of $A\mathbf{x} = \mathbf{b}$, em i.e., $T\mathbf{b} = \mathbf{x}$. It is easy to check that this is a linear transformation, since if $T\mathbf{b}_1 = \mathbf{x}_1$, i.e., $A\mathbf{x}_1 = \mathbf{b}_1$ and $T\mathbf{b}_2 = \mathbf{x}_2$, i.e., $A\mathbf{x}_2 = \mathbf{b}_2$, then $A(t_1\mathbf{x}_1 + t_2\mathbf{x}_2) = t_1A\mathbf{x}_1 + t_2A\mathbf{x}_2 = t_1\mathbf{b}_1 + t_2\mathbf{b}_2$, so that $T(t_1\mathbf{b}_1 + t_2\mathbf{b}_2) = t_1\mathbf{x}_1 + t_2\mathbf{x}_2 = t_1T\mathbf{b}_1 + t_2T\mathbf{b}_2$. Since T is a linear transformation, it is given by some matrix B , and since $T(A\mathbf{x}) = \mathbf{x}$, we must have $BA\mathbf{x} = \mathbf{x}$ which implies that BA is the identity matrix.

Going back to our first example, notice that not only is $BA = I$, but $AB = I$ too, since

$$\begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} = \begin{bmatrix} 6-5 & -2+2 \\ 15-15 & -5+6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

For a general choice of A and B , BA need not be equal to AB . But if B is the inverse of A , then it *is* always true that $AB = BA = I$.

To see this, suppose that A is invertible with $BA = I$ but we don't know yet whether $AB = I$. So what we need to show is that for every vector \mathbf{x} , $AB\mathbf{x} = \mathbf{x}$. First notice that if A is invertible, then any vector \mathbf{x} can be written in the form $A\mathbf{y}$ for some \mathbf{y} , since this is just the same as saying that the equation $A\mathbf{y} = \mathbf{x}$ has a solution \mathbf{y} . Thus $AB\mathbf{x} = AB A\mathbf{y} = A I \mathbf{y} = A\mathbf{y} = \mathbf{x}$.

4.5.1 Computing the inverse

How can we compute the inverse of an $n \times n$ matrix A ? Suppose that A is invertible and B is the inverse. Then $AB = I$. We can rewrite this equation as follows. Think of the columns of B as being column vectors so that

$$B = \left[\mathbf{b}_1 \mid \mathbf{b}_2 \mid \cdots \mid \mathbf{b}_n \right]$$

Then the rules of matrix multiplication imply that

$$AB = \left[A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_n \right]$$

Now the identity matrix can also be written as a matrix of column vectors. In this case the k th column is simply the matrix with zeros everywhere except for a 1 in the k th place, in other words the vector \mathbf{e}_k . Thus

$$I = \left[\mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n \right]$$

So if $AB = I$ then the n equations

$$\begin{aligned} A\mathbf{b}_1 &= \mathbf{e}_1 \\ A\mathbf{b}_2 &= \mathbf{e}_2 \\ &\vdots \\ A\mathbf{b}_n &= \mathbf{e}_n \end{aligned}$$

hold. If we solve each of these equations for $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$, then we have found the inverse B .

Here is a simple example. Suppose we want to find the inverse for

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}.$$

According to our discussion, we must solve $A\mathbf{b}_1 = \mathbf{e}_1$ and $A\mathbf{b}_2 = \mathbf{e}_2$. The augmented matrix for $A\mathbf{b}_1 = \mathbf{e}_1$ is

$$\left[\begin{array}{cc|c} 2 & 1 & 1 \\ 5 & 3 & 0 \end{array} \right]$$

We now perform a sequence of row operations. First divide the first row by 2. This gives

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

Now subtract 5 times the first row from the second row. This gives

$$\left[\begin{array}{cc|c} 1 & 1/2 & 1/2 \\ 0 & 1/2 & -5/2 \end{array} \right].$$

Now subtract the second row from the first row. This gives

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

Finally, multiply the second row by 2. This gives

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

Therefore $\mathbf{b}_1 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$. The augmented matrix for $A\mathbf{b}_2 = \mathbf{e}_2$ is

$$\left[\begin{array}{cc|c} 2 & 1 & 0 \\ 5 & 3 & 1 \end{array} \right]$$

We now perform a sequence of row operations. First divide the first row by 2. This gives

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

Now subtract 5 times the first row from the second row. This gives

$$\left[\begin{array}{cc|c} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \end{array} \right].$$

Now subtract the second row from the first row. This gives

$$\left[\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1/2 & 1 \end{array} \right].$$

Finally, multiply the second row by 2. This gives

$$\left[\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 2 \end{array} \right].$$

Therefore $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. So

$$B = [\mathbf{b}_1 | \mathbf{b}_2] = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$$

Notice that we performed *exactly the same sequence of row operations* in finding \mathbf{b}_1 and \mathbf{b}_2 . This is because the row operations only depend on the left side of the augmented matrix, in other words, the matrix A . If we used this procedure to find the inverse of an $n \times n$ matrix, we would end up doing exactly the same row operations n times. Clearly this is a big waste of effort! We can save a lot of work by solving all the equations at the same time. To do this we make a super-augmented matrix with both right sides.

$$\left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array} \right]$$

Now we only have to go through the sequence of row operations once, keeping track of both right sides simultaneously. Going through the same sequence, we obtain

$$\begin{aligned} & \left[\begin{array}{cc|cc} 1 & 1/2 & 1/2 & 0 \\ 5 & 3 & 0 & 1 \end{array} \right] \cdot \\ & \left[\begin{array}{cc|cc} 1 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & -5/2 & 1 \end{array} \right] \cdot \\ & \left[\begin{array}{cc|cc} 1 & 0 & 3 & -1 \\ 0 & 1/2 & -5/2 & 1 \end{array} \right] \cdot \\ & \left[\begin{array}{cc|cc} 1 & 0 & 3 & -1 \\ 0 & 1 & -5 & 2 \end{array} \right] \cdot \end{aligned}$$

Notice that the vectors \mathbf{b}_1 and \mathbf{b}_2 are automatically arranged as columns on the right side, so the matrix on the right is the inverse B .

The same procedure works for any size of square matrix. To find the inverse of A form the super-augmented matrix $[A|I]$. Then do a sequence of row operations to reduce A to the identity. If the resulting matrix is $[I|B]$ then B is the inverse matrix.

What happens if A doesn't have an inverse? In this case it will be impossible to reduce A to the identity matrix, since the rank of A is less than n . So the procedure will fail, as it must.

As another example, let us now compute the inverse of an arbitrary invertible 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

We will see that A is invertible precisely when its determinant $\Delta = ad - bc$ is non-zero. So let's assume this is the case, and do the computation. To start with, let's assume that $ac \neq 0$. Then neither a or c are zero. Here is the sequence of row transformations.

$$\left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right]$$

$$\left[\begin{array}{cc|cc} ac & bc & c & 0 \\ ac & ad & 0 & a \end{array} \right] \begin{array}{l} c(1) \\ a(2) \end{array}$$

Notice that multiplication by a and by c would not be legal row transformations if either a or c were zero.

$$\left[\begin{array}{cc|cc} ac & bc & c & 0 \\ 0 & ad - bc & -c & a \end{array} \right] (2) - (1)$$

$$\left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & 1 & -c/\Delta & a/\Delta \end{array} \right] \begin{array}{l} (1/ac)(1) \\ (1/\Delta)(2) \end{array}$$

$$\left[\begin{array}{cc|cc} 1 & 0 & d/\Delta & -b/\Delta \\ 0 & 1 & -c/\Delta & a/\Delta \end{array} \right] \begin{array}{l} (1) - (b/a)(2) \\ (1/\Delta)(2) \end{array}$$

Thus the inverse matrix is

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

This was derived under the additional assumption that $ac \neq 0$. However one can check directly that the same formula works, so long as $\Delta = ad - bc \neq 0$.

4.5.2 Inverses of Products

If both A and B are invertible, then so is AB . The inverse of AB is given by $B^{-1}A^{-1}$. To check this, simply compute

$$ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I.$$

If one of A or B is not invertible then AB is not invertible. To see this recall that a matrix C is not invertible exactly whenever there is a non-zero solution \mathbf{x} to $C\mathbf{x} = \mathbf{0}$. If B is not invertible, then there is a non-zero vector \mathbf{x} with $B\mathbf{x} = \mathbf{0}$. Then $AB\mathbf{x} = A\mathbf{0} = \mathbf{0}$ so AB is not invertible too. If B is invertible, but A is not, then there is a non-zero \mathbf{x} with $A\mathbf{x} = \mathbf{0}$. Let $\mathbf{y} = B^{-1}\mathbf{x}$. Since B^{-1} is invertible, \mathbf{y} cannot be zero. We have $AB\mathbf{y} = ABB^{-1}\mathbf{x} = A\mathbf{x} = \mathbf{0}$ so AB is not invertible.

4.5.3 MATLAB

If \mathbf{A} is an invertible $n \times n$ matrix then the MATLAB command `inv(A)` will return its inverse. A system of linear equations corresponding to $\mathbf{A}\mathbf{x} = \mathbf{b}$ where \mathbf{A} is an invertible $n \times n$ matrix and \mathbf{b} is a column vector with n components can be solved with the command

$$\mathbf{x} = \mathbf{A} \backslash \mathbf{b}$$

Note that if \mathbf{A} is not square or \mathbf{A} is square but not invertible the corresponding system cannot be solved in this way. In this case, use the MATLAB command `rref` on the augmented matrix of the system as described in Chapter 3.

4.5.4 Problems

Problem 4.30 Which of the following matrices are invertible?

$$(a) \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$(b) \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 1 & 1 \end{bmatrix}$$

$$(c) \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 1 & 2 \end{bmatrix}$$

Problem 4.31 Find the inverse for

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

Problem 4.32 Determine which of these matrices are invertible, and find the inverse for the invertible ones.

$$(a) \begin{bmatrix} 2 & 3 & -1 \\ 1 & 2 & 3 \\ -1 & -1 & 4 \end{bmatrix}$$

$$(b) \begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & -1 \\ 2 & -1 & 1 \end{bmatrix}$$

$$(c) \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$$

$$(d) \begin{bmatrix} 2 & 1 & 4 \\ 3 & 2 & 5 \\ 0 & -1 & 1 \end{bmatrix}$$

$$(e) \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$(f) \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$$

Problem 4.33 The following matrices are invertible. Find their inverses.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & -3 & 1 \\ -1 & 3 & -3 & -2 \\ 2 & 0 & 1 & 5 \\ 3 & 1 & -2 & 5 \end{bmatrix}.$$

Problem 4.34 Consider the system

$$\begin{cases} x + 5z = 6 \\ x - 2y + 3z = 14 \\ 2x + y - 3z = -2 \end{cases}$$

1. Write the above system in the matrix-form, in other words find a matrix

$$A \text{ such that } A\mathbf{x} = \begin{bmatrix} 6 \\ 14 \\ -2 \end{bmatrix}.$$

2. Find the inverse of A , if possible.
3. How many solutions the system have? Write them all down.

4.6 Determinants

4.6.1 Definition of Determinants

We have already encountered determinants for 2×2 and 3×3 matrices. For 2×2 matrices

$$\det \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

For 3×3 matrices we can define the determinant by expanding along the top row:

$$\det \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} = a_{1,1} \det \begin{bmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{bmatrix} - a_{1,2} \det \begin{bmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{bmatrix} + a_{1,3} \det \begin{bmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{bmatrix}$$

If we multiply out the 2×2 determinants in this definition we arrive at the expression

$$\det \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} = a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{2,3}a_{3,2} + a_{1,2}a_{2,3}a_{3,1} - a_{1,2}a_{2,1}a_{3,3} + a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{2,2}a_{3,1}$$

We now make a similar definition for an $n \times n$ matrix. Let A be an $n \times n$ matrix. Define $M_{i,j}$ to be the $(n-1) \times (n-1)$ matrix obtained by crossing out the i th row and the j th column. So, for example, if

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

then

$$M_{1,2} = \begin{bmatrix} \times & \times & \times & \times \\ 5 & \times & 7 & 8 \\ 9 & \times & 1 & 2 \\ 3 & \times & 5 & 6 \end{bmatrix} = \begin{bmatrix} 5 & 7 & 8 \\ 9 & 1 & 2 \\ 3 & 5 & 6 \end{bmatrix}$$

We now define the determinant of an $n \times n$ matrix A to be

$$\det(A) = a_{1,1} \det(M_{1,1}) - a_{1,2} \det(M_{1,2}) + \cdots \pm a_{1,n} \det(M_{1,n}) = \sum_{j=1}^n (-1)^{j+1} a_{1,j} \det(M_{1,j}).$$

Of course, this formula still contains determinants on the right hand side. However, they are determinants of $(n-1) \times (n-1)$ matrices. If we apply this definition to those determinants we get a more complicated formula involving $(n-2) \times (n-2)$ matrices, and so on, until we arrive at an extremely long expression (with $n!$ terms) involving only numbers.

Calculating an expression with $n!$ is completely impossible, even with the fastest computers, when n gets reasonable large. For example $100! \approx 10^{158}$. Yet, your computer at home can compute the determinant of a 100×100 matrix in less than a second. The secret, of course, is to compute the determinant in a different way. We start by computing the determinant of triangular matrices.

4.6.2 Determinants of Triangular matrices

Recall that triangular matrices are matrices whose entries above or below the diagonal are all zero. For 2×2 matrices

$$\det \begin{bmatrix} a_{1,1} & a_{1,2} \\ 0 & a_{2,2} \end{bmatrix} = a_{1,1}a_{2,2} - a_{1,2}0 = a_{1,1}a_{2,2}$$

and

$$\det \begin{bmatrix} a_{1,1} & 0 \\ a_{2,1} & a_{2,2} \end{bmatrix} = a_{1,1}a_{2,2} - 0a_{2,1} = a_{1,1}a_{2,2}$$

so the determinant is the product of the diagonal elements. For 3×3 matrices

$$\begin{aligned} \det \begin{bmatrix} a_{1,1} & 0 & 0 \\ a_{2,1} & a_{2,2} & 0 \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} &= a_{1,1} \det \begin{bmatrix} a_{2,2} & 0 \\ a_{3,2} & a_{3,3} \end{bmatrix} - 0 + 0 \\ &= a_{1,1}a_{2,2}a_{3,3} \end{aligned}$$

A similar expansion shows that the determinant of an $n \times n$ lower triangular matrix is the product of the diagonal elements. For upper triangular matrices we have

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

Since we already know that the determinant of a 2×2 triangular matrix is the product of the diagonals, we can see easily that the last two terms in this expression are zero. Thus we get

$$\begin{aligned} \det \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ 0 & a_{2,2} & a_{2,3} \\ 0 & 0 & a_{3,3} \end{bmatrix} &= a_{1,1} \det \begin{bmatrix} a_{2,2} & a_{2,3} \\ 0 & a_{3,3} \end{bmatrix} \\ &= a_{1,1} a_{2,2} a_{3,3} \end{aligned}$$

Once we know that the determinant of a 3×3 upper triangular matrix is the product of the diagonal elements, we can do a similar calculation to the one above to conclude that determinant of a 4×4 upper triangular matrix is the product of the diagonal elements, and so on.

Thus, the determinant of any (upper or lower) triangular $n \times n$ matrix is the product of the diagonal elements.

We know that an arbitrary $n \times n$ matrix can be reduced to an upper (or lower) triangular matrix by a sequence of row operations. This is the key to computing the determinant efficiently. We need to determine how the determinant of a matrix changes when we do an elementary row operation on it.

4.6.3 Summary of determinant calculation rules

We summarize the results on determinant calculation shown in more detail in the additional topics sections 4.7.7 to 4.7.13 below. The first three rules below show how a determinant changes when row operations are applied to it. This can be used to simplify determinant calculation as shown in the next section. In the statements below, A is a square matrix.

1. If B is obtained from A by multiplying *one* row of A by the constant c then $\det(B) = c \det(A)$.
2. If B is obtained from A by switching two rows of A then $\det(B) = -\det(A)$.
3. If B is obtained from A by adding a multiple of one row to another then $\det(B) = \det(A)$.
4. $\det(A) = 0$ if and only if A is not invertible.
5. For all square matrices B of the same size as A , $\det(AB) = \det(A) \det(B)$.
6. $\det(A^T) = \det(A)$.

4.6.4 Calculation of determinant using row operations

We can now use elementary row operations to compute the determinant of

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

The sequence of row operations that transforms this matrix into an upper triangular one is (R2)-(R1), (R3)-2(R1), exchange (R2) and (R3). The determinant doesn't change under the first two transformations, and changes sign under the third. Thus

$$\begin{aligned} \det\left(\begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & 3 & 0 \end{bmatrix}\right) &= \det\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -2 \\ 2 & 3 & 0 \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -2 \\ 0 & -1 & -6 \end{bmatrix}\right) \\ &= -\det\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -6 \\ 0 & 0 & -2 \end{bmatrix}\right) \\ &= -(1)(-1)(-2) = -2 \end{aligned}$$

4.6.5 More expansion formulae

We can use the properties of the determinant to derive alternative expansion formulae. Recall that we defined the determinant to be

$$\det(A) = \sum_{j=1}^n (-1)^{j+1} a_{1,j} \det(M_{1,j}).$$

In other words, we expanded along the top row. Now let's see that we can expand along other rows as well. Let A be the original matrix with rows $\mathbf{a}_1 = [a_{1,1}, a_{1,2}, \dots, a_{1,n}]$, \dots , $\mathbf{a}_n = [a_{n,1}, a_{n,2}, \dots, a_{n,n}]$. For example, if A is a 5×5 matrix then

$$A = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \mathbf{a}_4 \\ \mathbf{a}_5 \end{bmatrix}$$

Suppose we want to expand along the fourth row. Let A' be the matrix, where the fourth row of A has been moved to the first row, with all other rows still in the same order, i.e.,

$$A' = \begin{bmatrix} \mathbf{a}_4 \\ \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \mathbf{a}_5 \end{bmatrix}$$

How is the determinant of A' related to the determinant of A ? We can change A to A' be a series of row flips as follows:

$$A = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \mathbf{a}_4 \\ \mathbf{a}_5 \end{bmatrix}, \quad \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_4 \\ \mathbf{a}_3 \\ \mathbf{a}_5 \end{bmatrix}, \quad \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_4 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \mathbf{a}_5 \end{bmatrix}, \quad \begin{bmatrix} \mathbf{a}_4 \\ \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \mathbf{a}_5 \end{bmatrix} = A'$$

We have performed 3 flips, so $\det(A') = (-1)^3 \det(A) = -\det(A)$.

In general, to move the i th row to the top in this way, we must perform $i - 1$ flips, so $\det(A') = (-1)^{i-1} \det(A)$

Notice that A' is a matrix with the properties

- (1) $a'_{1,j} = a_{i,j}$, since we have moved the i th row to the top
- (2) $M'_{1,j} = M_{i,j}$, since we haven't changed the order of the other rows.

Therefore

$$\begin{aligned} \det(A) &= (-1)^{i-1} \det(A') \\ &= (-1)^{i-1} \sum_{j=1}^n (-1)^{j+1} a'_{1,j} \det(M'_{1,j}) \\ &= (-1)^{i-1} \sum_{j=1}^n (-1)^{j+1} a_{i,j} \det(M_{i,j}) \\ &= \sum_{j=1}^n (-1)^{i+j} a_{i,j} \det(M_{i,j}) \end{aligned}$$

This is the formula for expansion along the i th row.

As an example let's compute the determinant of a 3×3 matrix by expanding along the second row.

$$\begin{aligned} \det \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 2 & 1 \end{bmatrix} &= -\det \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} + 3 \det \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} - \det \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \\ &= -2 + 6 + 3 - 9 - 2 + 2 = -2 \end{aligned}$$

The formula for expanding along the i th row is handy if the matrix happens to have a row with many zeros.

Using the fact that $\det(A) = \det(A^T)$ we can also write down expansion formulae along columns, since the columns of A are the rows of A^T . We end up with the formula

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{i,j} \det(M_{i,j})$$

As an example let's compute the determinant of a 3×3 matrix by expanding along the second column.

$$\begin{aligned} \det \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 2 & 1 \end{bmatrix} &= -2 \det \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + 3 \det \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} - 2 \det \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix} \\ &= -2 + 2 + 3 - 9 - 2 + 6 = -2 \end{aligned}$$

The formula for expanding along the j th column is handy if the matrix happens to have a column with many zeros.

4.6.6 MATLAB

The MATLAB command `det(A)` can be used to determine the determinant of a square matrix as introduced in Chapter 3.

4.6.7 Problems

Problem 4.35 Find the determinant of

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix}$$

Problem 4.36 Find the determinant of

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 2 & 4 & 8 \\ 1 & -2 & 4 & -8 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Problem 4.37 Consider the matrix

$$A = \begin{bmatrix} 2 & 0 & 2 & 4 \\ 0 & 0 & 3 & 2 \\ 2 & 2 & 4 & 4 \\ 3 & 0 & 6 & 2 \end{bmatrix}.$$

1. Find the determinant of A by expanding on the first row.
2. Find the determinant of A by expanding on the row or column of your choice. This expansion should be simpler than the one used in part (a).

Problem 4.38 Compute

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

by expanding along the second row, and by expanding along the third column.

Problem 4.39 Find the determinant of A given above by using row operations to put A into echelon form.

Problem 4.40 We know that the determinant of the $n \times n$ matrix

$$\begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix}$$

is $a_1 a_2 \cdots a_n$ if the blank entries represent 0. What is the equivalent formula for the determinant of the $n \times n$ matrix

$$\begin{bmatrix} & & & a_n \\ & & \ddots & \\ & a_2 & & \\ a_1 & & & \end{bmatrix} ?$$

Justify your answer.

Problem 4.41 Find all values of λ for which the matrix

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

is not invertible.

Problem 4.42 Determine whether each of the following statements is true or false. If true, explain briefly why it is true without showing specific examples. If false, explain briefly why it is false or give a simple counterexample, and give a correct statement by making small changes to the underlined text.

1. The determinant of a square matrix is the product of the entries on its main diagonal.
2. For every matrix A , we have $\det(AA^T) = \det(A^T A)$
3. If A and B are both invertible $n \times n$ matrices, then $\det(B^{-1}AB) = \det(BAB^{-1}) = \det(A)$.
4. If an $n \times n$ matrix A is multiplied by a scalar c , the determinant of the resulting matrix is $c \cdot \det(A)$.
5. If the column vectors of an $n \times n$ matrix are linearly independent, then the row vectors of the same matrix are also linearly independent.

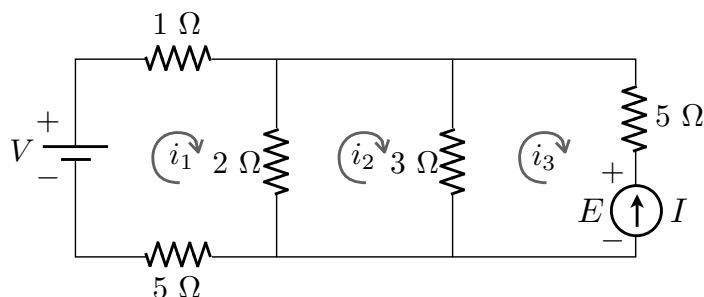


Figure 4.10: The circuit considered in Example 3.15.

4.7 Additional Topics

4.7.1 Return to Resistor Networks

Let us revisit the resistor networks we considered in section 3.5. These networks consist of resistors and voltage and current sources. The *fundamental problem* was introduced, which was to write the current through the voltage sources and the voltage across current sources in terms of arbitrary values of the sources.

Recall example 3.15 which considered the circuit shown again in figure 4.10. The sources are V and I . The *fundamental problem* is to determine E and J (the current through the voltage source) in terms of V and I . The solution found in section 3.5.3 was

$$\begin{aligned} E &= 6I + V/6 \\ J &= -\frac{1}{6}I + \frac{10}{72}V \end{aligned} \quad (4.5)$$

This solution can be given in matrix-vector form as

$$\begin{bmatrix} E \\ J \end{bmatrix} = \begin{bmatrix} 6 & \frac{1}{6} \\ -\frac{1}{6} & \frac{10}{72} \end{bmatrix} \begin{bmatrix} I \\ V \end{bmatrix} \quad (4.6)$$

where we will call the 2×2 matrix in the equation above F . It can be shown that for any circuit with n sources, the solution of the fundamental problem can be written as multiplication by an $n \times n$ matrix. We will show below how F can be constructed in a systematic way. Look back to section 3.5.3 to see how we arrived at (4.5). There were three parts to the process:

1. We wrote the linear system for the loop currents and the voltage drop across the current source. The equations for the system had the sources I and V in the right hand sides.
2. We solved the system symbolically in terms of the right hand side involving V and I .
3. We identified the elements of the solution that solved the fundamental problem.

We will now describe these three steps by matrix multiplication and do the multiplication as an alternate way to determine the matrix F in (4.6). Recall that the unknowns for the circuit are the three loop currents and E , the voltage across the current source. Define the intermediate unknown column vector

$$\mathbf{x} = (i_1, i_2, i_3, E)^T.$$

Proceeding as in example 3.15 we find equations for \mathbf{x} by considering the voltage drops around each elementary loop (the first three equations) and matching the current through the current source (the last equation below):

$$\begin{array}{ccccccc} 8i_1 & -2i_2 & & & = & V \\ -2i_1 & +5i_2 & -3i_3 & & = & 0 \\ & -3i_2 & +8i_3 & +E & = & 0 \\ & & i_3 & & = & -I \end{array} \quad (4.7)$$

Here, we did not eliminate i_3 from the equations as was done in section 3.5.3. The system (4.7) can be written

$$F_2 \mathbf{x} = F_1 \begin{bmatrix} I \\ V \end{bmatrix} \quad (4.8)$$

where

$$F_2 = \begin{bmatrix} 8 & -2 & 0 & 0 \\ -2 & 5 & -3 & 0 \\ 0 & -3 & 8 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad F_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ -1 & 0 \end{bmatrix}$$

Since F_2 is invertible (always true for our approach using loop currents and current source voltages as variables) we can proceed from (4.8) to

$$\mathbf{x} = F_2^{-1} F_1 \begin{bmatrix} I \\ V \end{bmatrix} \quad (4.9)$$

Now to solve the *fundamental problem* of the circuit we want $J = i_1$, the current through the voltage sources, and E , the voltage across the current source. We can write

$$\begin{bmatrix} E \\ J \end{bmatrix} = F_3 \mathbf{x} \quad (4.10)$$

where

$$F_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Combining (4.9) and (4.10) we obtain

$$\begin{bmatrix} E \\ J \end{bmatrix} = F_3 F_2^{-1} F_1 \begin{bmatrix} I \\ V \end{bmatrix} \quad (4.11)$$

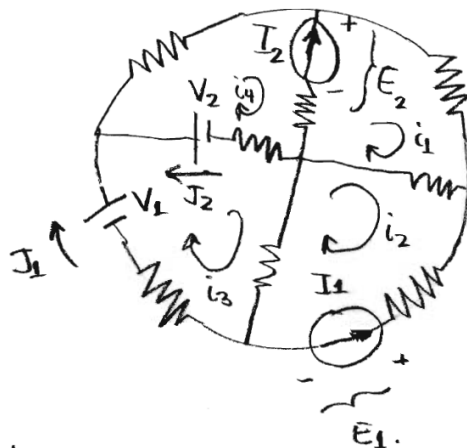


Figure 4.11: The circuit considered in Example 4.3. All resistors are 1Ω .

Comparing (4.11) to (4.6) we see that

$$F = F_3 F_2^{-1} F_1$$

where F is the matrix representing the fundamental solution. Computation does indeed show that $F_3 F_2^{-1} F_1$ equals F with the matrices defined above.

The solution of the fundamental problem for any circuit can be written as the product of three matrices in the same process described above. We illustrate this with a more complex circuit below.

Example 4.3 Find the matrix for the fundamental solution of the circuit shown in Figure 4.11. Here there are four sources I_1 , I_2 , V_1 and V_2 so the fundamental problem will be written

$$\begin{bmatrix} E_1 \\ E_2 \\ J_1 \\ J_2 \end{bmatrix} = F \begin{bmatrix} I_1 \\ I_2 \\ V_1 \\ V_2 \end{bmatrix} \quad (4.12)$$

where F is a 4×4 matrix to be determined. The intermediate unknowns in the circuit will be

$$\mathbf{x} = (i_1, i_2, i_3, i_4, E_1, E_2)^T$$

as shown in the figure. Equations for these unknowns are found using the loop current method:

$$\begin{array}{rcccccc} 3i_1 & -i_2 & & -i_4 & & -E_2 & = & 0 \\ -i_1 & +3i_2 & -i_3 & & +E_1 & & = & 0 \\ & -i_2 & +3i_3 & -i_4 & & & = & V_1 - V_2 \\ -i_1 & & -i_3 & +3i_4 & & +E_2 & = & V_1 \\ & i_2 & & & & & = & -I_1 \\ -i_1 & & & +i_4 & & & = & -I_2 \end{array} \quad (4.13)$$

where the first four equations above come from matching voltage drops around the four elementary loops and the last two equations come from matching the loop currents to the current sources. Note that (4.13) can be written

$$F_2 \mathbf{x} = F_1 \begin{bmatrix} I_1 \\ I_2 \\ V_1 \\ V_2 \end{bmatrix} \quad (4.14)$$

where

$$F_2 = \begin{bmatrix} 3 & -1 & 0 & -1 & 0 & -1 \\ -1 & 3 & -1 & 0 & 1 & 0 \\ 0 & -1 & 3 & -1 & 0 & 0 \\ -1 & 0 & -1 & 3 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad F_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

Now $J_2 = i_4 - i_3$ and $J_1 = i_3$ so we can write

$$\begin{bmatrix} E_1 \\ E_2 \\ J_1 \\ J_2 \end{bmatrix} = F_3 \mathbf{x} \quad (4.15)$$

where

$$F_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \end{bmatrix}.$$

As above we can combine (4.14) and (4.15) to obtain

$$\begin{bmatrix} E_1 \\ E_2 \\ J_1 \\ J_2 \end{bmatrix} = F_3 F_2^{-1} F_1 \begin{bmatrix} I_1 \\ I_2 \\ V_1 \\ V_2 \end{bmatrix}.$$

Comparing the equation above to the desired form (4.12) we see that the fundamental matrix F can be computed as

$$F = F_3 F_2^{-1} F_1.$$

A MATLAB computation with the matrices defined above gives

$$F \approx \begin{bmatrix} 2.1818 & 0.2727 & 0.8182 & -0.4545 \\ 0.2727 & 1.9091 & 0.7273 & -0.1818 \\ -0.4545 & -0.1818 & 0.4545 & -0.3636 \\ 0.0909 & -0.3636 & -0.0909 & 0.2727 \end{bmatrix}$$

This technique can be used to find the fundamental solution of circuits even if they are very large. You can imagine that the process of setting up the matrices F_1 , F_2 and F_3 , computing an approximation of F_2^{-1} and multiplying the matrices together can all be automated and done computationally.

In Chapter 5, fundamental matrices for circuits will be used to investigate the transient behaviour of circuits with capacitors and inductors as well as resistors.

4.7.2 Application: General Least Squares

Let us restate our results from Chapter 3 on minimization of quadratic functions using matrix notation. A quadratic function of

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

in \mathbf{R}^n can be written in matrix form as

$$f(\mathbf{x}) = \mathbf{x} \cdot A\mathbf{x} + \mathbf{b} \cdot \mathbf{x} + c$$

where A is an $n \times n$ matrix and $\mathbf{b} \in \mathbf{R}^n$ and c is a number. The vector containing the partial derivatives can be computed to be

$$\begin{bmatrix} \partial f / \partial x_1 \\ \vdots \\ \partial f / \partial x_n \end{bmatrix} = (A + A^T)\mathbf{x} + \mathbf{b}$$

Recall that we made the assumption that $a_{ij} = a_{ji}$ when we considered this problem before. This property can be stated in compact form as $A = A^T$. If this is true then $(A + A^T) = 2A$ so

$$\begin{bmatrix} \partial f / \partial x_1 \\ \vdots \\ \partial f / \partial x_n \end{bmatrix} = 2A\mathbf{x} + \mathbf{b}$$

To find the minimum value of f (if it exists) we need to find the value of \mathbf{x} for which the vector above is zero. In other words, \mathbf{x} solves the equation

$$2A\mathbf{x} = -\mathbf{b}.$$

This is the same equation that we derived before.

4.7.3 Least squares solutions

Let's take another look the situation where a system of linear equations, which we now can write

$$B\mathbf{x} = \mathbf{c},$$

has no solution. Typically this will be the case if there are more equations than variables, that is, B is an matrix with more rows than columns. In this case there is no value of \mathbf{x} that makes the left side equal the right side. However, we may try to find the value of \mathbf{x} for which the right side $B\mathbf{x}$ is *closest* to the left side \mathbf{c} .

One way to go about this is to try to minimize distance between the left and right sides. It is more convenient to minimize the square of the distance. This quantity can be written

$$\begin{aligned}\|B\mathbf{x} - \mathbf{c}\|^2 &= (B\mathbf{x} - \mathbf{c}) \cdot (B\mathbf{x} - \mathbf{c}) \\ &= (B\mathbf{x}) \cdot (B\mathbf{x}) - (B\mathbf{x}) \cdot \mathbf{c} - \mathbf{c} \cdot (B\mathbf{x}) + \mathbf{c} \cdot \mathbf{c} \\ &= \mathbf{x} \cdot (B^T B\mathbf{x}) - 2(B^T \mathbf{c}) \cdot \mathbf{x} + \mathbf{c} \cdot \mathbf{c}\end{aligned}$$

This is a quadratic function, written in matrix form. We want to use the formula of the previous section with $A = B^T B$ and $\mathbf{b} = B^T \mathbf{c}$. Before we can do so, we must verify that $A = B^T B$ satisfies $A^T = A$. This is true because

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

Thus the formula of the previous section implies that the minimum occurs at the value of \mathbf{x} that solves the linear equation

$$B^T B\mathbf{x} = B^T \mathbf{c}$$

Here we have cancelled a factor of 2 on each side.

Now let's derive the same result in another way. Think of all the values of $B\mathbf{x}$, as \mathbf{x} ranges through all possible values in \mathbf{R}^n as forming a (high dimensional) plane in \mathbf{R}^m . Our goal is to find the value of \mathbf{x} so that the corresponding value of $B\mathbf{x}$ on the plane is closest to \mathbf{c} . Using the analogy to the geometric picture in three dimensions, we see that the minimum will occur when $B\mathbf{x} - \mathbf{c}$ is orthogonal to the plane. This means that the dot product of $B\mathbf{x} - \mathbf{c}$ with every vector in the plane, that is, every vector of the form $B\mathbf{y}$, should be zero. Thus we have

$$(B\mathbf{y}) \cdot (B\mathbf{x} - \mathbf{c}) = 0$$

for every $\mathbf{y} \in \mathbf{R}^n$. This is the same as

$$\mathbf{y} \cdot (B^T (B\mathbf{x} - \mathbf{c})) = \mathbf{y} \cdot (B^T B\mathbf{x} - B^T \mathbf{c}) = 0$$

for every $\mathbf{y} \in \mathbf{R}^n$. This can happen only if

$$B^T B\mathbf{x} = B^T \mathbf{c}$$

which is the same result we obtained before.

4.7.4 Problems

Problem 4.43 Find the least squares solution to

$$\begin{array}{rcl}x_1 & +x_2 & = 1 \\x_1 & & = 1 \\x_1 & +x_2 & = 0\end{array}$$

Compare $B\mathbf{x}$ and \mathbf{b} .

Problem 4.44 Refer back to the least squares fit example, where we tried to find the best straight line going through a collection of points (x_i, y_i) . Another way of formulating this problem is this. The line $y = ax + b$ passes through the point (x_i, y_i) if

$$ax_i + b = y_i \quad (4.16)$$

So, saying that the straight line passes through all n points is the same as saying that a and b solve the system of n linear equations given by (4.16) for $i = 1, \dots, n$. Of course, unless the points all actually lie on the same line, this system of equations has no solutions. Show that the least squares solution to this problem is the same as we obtained before. (You may simplify the problem by assuming there are only three points (x_1, y_1) , (x_2, y_2) and (x_3, y_3) .)

4.7.5 Elementary matrices

Recall that there are three row operation that are used in Gaussian elimination: (1) multiplication of a row by a non-zero number, (2) add a multiple of one row to another row and (3) exchanging two rows.

It turns out that each elementary row operation can be implemented by left multiplication by a matrix. In other words, for each elementary row operation there is a matrix Q such that QA is what you get by doing that row operation to the matrix A .

Here is an example. Suppose

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

and suppose that the row operation is multiplying the first row by 2. Then the matrix you get by doing that row operation to the matrix A is

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

In this case the matrix Q turns out to be

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

Since

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

i.e., $QA = A'$.

Now suppose that the elementary row operation is subtracting twice the first row from the second row. Then the matrix you get by doing that row operation to the matrix A is

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

In this case the matrix Q turns out to be

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

Since

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

i.e., again, $QA = A'$.

Finally, suppose that the elementary row operation is exchanging the second and the third rows. Then the matrix you get by doing that row operation to the matrix A is

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

In this case the matrix Q turns out to be

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

Since

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

i.e., again, $QA = A'$.

How can we find the matrices Q (called elementary matrices)? Here is the procedure. Start with the identity matrix I and do the row transformation to it. The resulting matrix Q is the matrix that implements that row transformation by multiplication from the left. Notice that this is true in the examples above. In the first example, the row transformation was multiplying the first row by 2. If you multiply the first row of

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

by two you get

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

In the second example, the row transformation was subtracting twice the first row from the second row. If you subtract twice the second row from the first row of I by two you get

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

In the third example, the row transformation was exchanging the second and third rows. If you exchange the second and third rows of I , you get

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

Elementary matrices are useful in theoretical studies of the Gaussian elimination process. We will use them briefly when studying determinants.

Suppose A is a matrix, and R is its reduced form. Then we can obtain R from A via a sequence of elementary row operations. Suppose that the corresponding elementary matrices are Q_1, Q_2, \dots, Q_k . Then, starting with A , the matrix after the first elementary row operation is Q_1A , then after the second elementary row operation is Q_2Q_1A , and so on, until we have

$$Q_kQ_{k-1}\cdots Q_2Q_1A = R.$$

Now let us apply the inverse matrices, starting with Q_k^{-1} . This gives

$$Q_k^{-1}Q_kQ_{k-1}\cdots Q_2Q_1A = Q_{k-1}\cdots Q_2Q_1A = Q_k^{-1}R.$$

Continuing in this way we see that

$$A = Q_1^{-1}Q_2^{-1}\cdots Q_k^{-1}R$$

In the special case that A is an $n \times n$ invertible matrix, A can be reduced to the identity matrix. In other words, we can take $R = I$. In this case A can be written as a product of elementary matrices.

$$A = Q_1^{-1}Q_2^{-1}\cdots Q_k^{-1}I = Q_1^{-1}Q_2^{-1}\cdots Q_k^{-1}$$

Notice that in this case

$$A^{-1} = Q_kQ_{k-1}\cdots Q_2Q_1.$$

As an example, let us write the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$$

as a product of elementary matrices. The sequence of row transformations that reduce A to the identity are:

- 1) $(1/2)(R1)$
- 2) $(R2)-5(R1)$
- 3) $(R1)-(R2)$
- 4) $2(R2)$

The corresponding elementary matrices and their inverses are

$$Q_1 = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} \quad Q_1^{-1} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Q_2 = \begin{bmatrix} 1 & 0 \\ -5 & 1 \end{bmatrix} \quad Q_2^{-1} = \begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix}$$

$$Q_3 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \quad Q_3^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$Q_4 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad Q_4^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$$

Therefore

$$A = Q_1^{-1}Q_2^{-1}Q_3^{-1}Q_4^{-1}$$

or

$$\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$$

4.7.6 Problems

Problem 4.45 *Each elementary matrix is invertible, and the inverse is also an elementary matrix. Find the inverses of the three examples of elementary matrices above. Notice that the inverse elementary matrix is the matrix for the row transformation that undoes the original row transformation.*

Problem 4.46 *Write the matrix*

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

as a product of elementary matrices.

4.7.7 Exchanging two rows changes the sign of the determinant

We start with the elementary row operation of exchanging two rows. For 2×2 determinants,

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc,$$

while

$$\det \begin{bmatrix} c & d \\ a & b \end{bmatrix} = cb - da = -(ad - bc),$$

so exchanging two rows changes the sign of the determinant.

We can do a similar calculation for 3×3 matrices. Its a bit messier, but still manageable. Again, we find that exchanging two rows changes the sign of the determinant.

How about the $n \times n$ case? We will assume that we have already proved the result for the $(n-1) \times (n-1)$ case, and show how we can use this to show the result for an $n \times n$ matrix. Thus knowing the result for 2×2 matrices, implies it for 3×3 , which in turn implies it for 4×4 matrices, and so on. We consider three cases, depending on which rows we are exchanging. Suppose A is the original matrix and A' is the matrix with two rows exchanged.

- (1) Exchanging two rows other than the first row: In this case we cross out the first row and any column from A' we obtain $M'_{1,j}$ which is the same as the matrix $M_{1,j}$ (corresponding to A) except with two of its rows exchanged. Since the size of $M_{1,j}$ is $n-1$ we know that $\det(M'_{1,j}) = -\det(M_{1,j})$ so

$$\begin{aligned} \det(A') &= \sum_{j=1}^n (-1)^{j+1} a_{1,j} \det(M'_{1,j}) \\ &= -\sum_{j=1}^n (-1)^{j+1} a_{1,j} \det(M_{1,j}) \\ &= -\det(A) \end{aligned}$$

- (2) Exchanging the first and second row. Do see that this changes the sign of the determinant we have to expand the expansion. The following is a bit sketchy. I'll probably skip it in class, but give the argument here for completeness. If we expand $M_{1,j}$ we get

$$\det(M_{1,j}) = \sum_{k=1}^{j-1} (-1)^{k+1} a_{2,k} \det(M_{1,2,j,k}) + \sum_{k=j+1}^n (-1)^k a_{2,k} \det(M_{1,2,j,k})$$

where $M_{1,2,j,k}$ is the matrix obtained from A by deleting the first and second rows, and the j th and k th columns. Inserting this into the expansion for A gives

$$\det(A) = \sum_{j=1}^n \sum_{k=1}^{j-1} (-1)^{j+k} a_{1,j} a_{2,k} \det(M_{1,2,j,k}) - \sum_{j=1}^n \sum_{k=j+1}^n (-1)^{j+k} a_{1,j} a_{2,k} \det(M_{1,2,j,k})$$

The sum splits into two parts. Flipping the first two rows of A just exchanges the two sums. In other words $S - R$ becomes $R - S$ which is $-(S - R)$. So exchanging the first two rows also changes the sign of the determinant.

- (3) Exchanging the first row with the k th row. We can effect this exchange by first exchanging the k th and the second row, then exchanging the first and the second row, then exchanging the k th and the second row again. Each flip changes the determinant by a minus sign, and since there are three flips, the overall change is by a minus sign.

Thus we can say that for any $n \times n$ matrix, exchanging two rows changes the sign of the determinant.

One immediate consequence of this fact is that a matrix with two rows the same has determinant zero. This is because if exchange the two rows the determinant changes by a minus sign, but the matrix doesn't change. Thus $\det(A) = -\det(A)$ which is only possible if $\det(A) = 0$.

4.7.8 The determinant is linear in each row separately

To say that the determinant is linear in the j th row means that if we write a matrix as a matrix of row vectors,

$$A = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_j \\ \vdots \\ \mathbf{a}_n \end{bmatrix}$$

then

$$\det \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ s\mathbf{b} + t\mathbf{c} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = s \det \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{b} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + t \det \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{c} \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

It is easy to from the expansion formula that the determinant is linear in the first row. For a 3×3 example we have

$$\begin{aligned} & \det \begin{pmatrix} sb_1 + tc_1 & sb_2 + tc_2 & sb_3 + tc_3 \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \\ &= (sb_1 + tc_1) \det(M_{1,1}) - (sb_2 + tc_2) \det(M_{1,2}) + (sb_3 + tc_3) \det(M_{1,3}) \\ &= s(b_1 \det(M_{1,1}) - b_2 \det(M_{1,2}) + b_3 \det(M_{1,3})) \\ &\quad + t(c_1 \det(M_{1,1}) - c_2 \det(M_{1,2}) + c_3 \det(M_{1,3})) \\ &= s \det \begin{pmatrix} b_1 & b_2 & b_3 \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} + t \det \begin{pmatrix} c_1 & c_2 & c_3 \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \end{aligned}$$

A similar calculation can be done for any $n \times n$ matrix to show linearity in the first row. To show linearity in some other row, we first swap that row and the first row, then use linearity in the first row, and then swap back again. So

$$\begin{aligned} \det \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ s\mathbf{b} + t\mathbf{c} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} &= -\det \begin{pmatrix} s\mathbf{b} + t\mathbf{c} \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \\ &= -s \det \begin{pmatrix} \mathbf{b} \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} - t \det \begin{pmatrix} \mathbf{c} \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \\ &= s \det \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{b} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + t \det \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{c} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \end{aligned}$$

Notice that linearity in each row separately does **not** mean that $\det(A + B) = \det(A) + \det(B)$.

Note that multiplying a row by a constant multiplies the determinant by the constant. This is a special case of linearity.

4.7.9 Adding a multiple of one row to another doesn't change the determinant

Now we will see that the most often used row operation—adding a multiple of one row to another—doesn't change the determinant at all. Let A be an $n \times n$ matrix. Write A as a matrix of rows.

$$A = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_j \\ \vdots \\ \mathbf{a}_n \end{bmatrix}$$

Adding s times the i th row to the j th row yields

$$A' = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_j + s\mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{bmatrix}$$

So

$$\det(A') = \det\left(\begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_j + s\mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{bmatrix}\right) = \det\left(\begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_j \\ \vdots \\ \mathbf{a}_n \end{bmatrix}\right) + s \det\left(\begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{bmatrix}\right) = \det(A) + 0$$

Here we used linearity in a row and the fact that the determinant of a matrix with two rows the same is zero.

4.7.10 The determinant of QA

To begin, we compute the determinants of the elementary matrices. Recall that if A' is the matrix obtained from A by an elementary row operation, then

- (1) $\det(A') = -\det(A)$ if the row operation is swapping two rows
- (2) $\det(A') = s \det(A)$ if the row operation is multiplying a row by s
- (3) $\det(A') = \det(A)$ if the row operation is adding a multiple of one row to another

Recall that the elementary matrices are obtained from the identity matrix I by an elementary row operation. So we can take $A = I$ and $A' = Q$ in the formulae above to obtain

- (1) $\det(Q) = -\det(I) = -1$ if the row operation is swapping two rows
- (2) $\det(Q) = s \det(I) = s$ if the row operation is multiplying a row by s
- (3) $\det(Q) = \det(I) = 1$ if the row operation is adding a multiple of one row to another

Going back to the first set of formulae, we have that in each case $A' = QA$. In each case the factor in front of $\det(A)$ is exactly $\det(Q)$. So we see that in each case

$$\det(QA) = \det(Q) \det(A).$$

This formula can be generalized. If Q_1, Q_2, \dots, Q_k are elementary matrices then $\det(Q_1 Q_2 Q_3 \cdots Q_k A) = \det(Q_1) \det(Q_2 Q_3 \cdots Q_k A) = \det(Q_1) \det(Q_2) \det(Q_3 \cdots Q_k A)$ and so on, so we arrive at the formula

$$\det(Q_1 Q_2 Q_3 \cdots Q_k A) = \det(Q_1) \det(Q_2) \cdots \det(Q_k) \det(A).$$

4.7.11 The determinant of A is zero exactly when A is not invertible

Recall that if R denotes the reduced form of A , obtained by performing the sequence of row reductions corresponding to Q_1, Q_2, \dots, Q_k , then

$$A = Q_1^{-1} Q_2^{-1} \cdots Q_k^{-1} R$$

Each Q_i^{-1} is an elementary matrix, therefore

$$\det(A) = \det(Q_1^{-1}) \det(Q_2^{-1}) \cdots \det(Q_k^{-1}) \det(R)$$

If A is not invertible, then R has a row of zeros along the bottom. Thus R is an upper triangular matrix with at least one zero on the diagonal. The determinant of R is the product of the diagonal elements so $\det(R) = 0$. Thus $\det(A) = 0$ too.

If A is invertible, then we can reduce A to the identity matrix. In other words, we can take $R = I$. Then $\det(R) = 1$. Each $\det(Q_i^{-1})$ is non-zero too, so $\det(A) \neq 0$.

4.7.12 The product formula: $\det(AB) = \det(A) \det(B)$

If either A or B is non-invertible, then AB is non-invertible too. Thus $\det(AB) = 0$ and one of $\det(A)$ or $\det(B)$ is zero, so $\det(A) \det(B) = 0$ too. Thus $\det(AB) = \det(A) \det(B)$.

If both A and B are invertible, then

$$A = Q_1^{-1} Q_2^{-1} \cdots Q_k^{-1}$$

so

$$\det(A) = \det(Q_1^{-1}) \det(Q_2^{-1}) \cdots \det(Q_k^{-1})$$

and

$$B = \tilde{Q}_1^{-1} \tilde{Q}_2^{-1} \cdots \tilde{Q}_j^{-1}$$

so

$$\det(B) = \det(\tilde{Q}_1^{-1}) \det(\tilde{Q}_2^{-1}) \cdots \det(\tilde{Q}_j^{-1})$$

Therefore

$$AB = Q_1^{-1}Q_2^{-1} \cdots Q_k^{-1}\tilde{Q}_1^{-1}\tilde{Q}_2^{-1} \cdots \tilde{Q}_j^{-1}$$

so

$$\det(AB) = \det(Q_1^{-1}) \det(Q_2^{-1}) \cdots \det(Q_k^{-1}) \det(\tilde{Q}_1^{-1}) \det(\tilde{Q}_2^{-1}) \cdots \det(\tilde{Q}_j^{-1}) = \det(A) \det(B)$$

4.7.13 The determinant of the transpose

Recall that the transpose A^T of a matrix A is the matrix you get when you flip A about its diagonal. If A is an $n \times n$ matrix, so is A^T and we can ask what the relationship between the determinants of these two matrices is. It turns out that they are the same.

$$\det(A^T) = \det(A).$$

If A is an upper or lower triangular matrix, this follows from the fact that the determinant of a triangular matrix is the product of the diagonal entries. If A is an arbitrary $n \times n$ matrix then the formula follows from two facts.

- (1) The transpose of a product of two matrices is given by $(AB)^T = B^T A^T$. This implies that $(A_1 A_2 \cdots A_n)^T = A_n^T \cdots A_2^T A_1^T$.
- (2) For an elementary matrix Q we have $\det(Q^T) = \det(Q)$.

If you accept these two facts, then we may write

$$A = Q_1^{-1}Q_2^{-1} \cdots Q_k^{-1}R$$

where R is upper triangular. Thus

$$A^T = R^T(Q_k^{-1})^T \cdots (Q_2^{-1})^T(Q_1^{-1})^T$$

so

$$\begin{aligned} \det(A^T) &= \det(R^T) \det((Q_k^{-1})^T) \cdots \det((Q_1^{-1})^T) \\ &= \det(R) \det(Q_k^{-1}) \cdots \det(Q_1^{-1}) \\ &= \det(Q_1^{-1}) \cdots \det(Q_k^{-1}) \det(R) \\ &= \det(A) \end{aligned}$$

4.7.14 An impractical formula for the inverse

We can use the expansion formulae of the previous section to obtain a formula for the inverse of a matrix A . This formula is really only practical for 3×3 matrices, since for larger matrices, the work involved in computing the determinants appearing is prohibitive.

We begin with the expansion formula

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{i,j} \det(M_{i,j})$$

If A is invertible, then $\det(A) \neq 0$ so we can divide by it to obtain

$$1 = \sum_{j=1}^n a_{i,j} \frac{(-1)^{i+j} \det(M_{i,j})}{\det(A)}$$

Now suppose we take the matrix A and replace the i th row by the k th row for some $k \neq i$. The resulting matrix A' has two rows the same, so its determinant is zero. Its expansion is the same as that for A , except that $a_{i,j}$ is replaced by $a_{k,j}$. Thus, if $k \neq i$

$$0 = \sum_{j=1}^n (-1)^{i+j} a_{k,j} \det(M_{i,j})$$

Dividing by $\det(A)$ yields

$$0 = \sum_{j=1}^n a_{k,j} \frac{(-1)^{i+j} \det(M_{i,j})}{\det(A)}$$

Now let B be the matrix with entries

$$b_{i,j} = \frac{(-1)^{i+j} \det(M_{j,i})}{\det(A)}$$

This turns out to be the inverse A^{-1} .

It gets a bit confusing with all the indices, but let's think about what we need to show. The k, i th entry of the product AB is given by

$$\begin{aligned} (AB)_{k,i} &= \sum_{j=1}^n a_{k,j} b_{j,i} \\ &= \sum_{j=1}^n a_{k,j} \frac{(-1)^{i+j} \det(M_{i,j})}{\det(A)} \end{aligned}$$

According to the formulae above, this sum is equal to 1 if $k = i$ and equal to 0 if $k \neq i$. In other words, AB is the identity matrix I . This shows that $B = A^{-1}$. Remember that

4.7.15 Cramer's rule, an impractical way to solve systems

Given an impractical way to compute the inverse, we can derive an impractical formula for the solution of a system of n equations in n unknowns, i.e., a matrix equation

$$A\mathbf{x} = \mathbf{b}$$

The solution \mathbf{x} is equal to $A^{-1}\mathbf{b}$. So if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, then using

the formula of the previous section for the inverse, we get

$$x_i = \sum_{j=1}^n A_{i,j}^{-1} b_j = \sum_{j=1}^n \frac{(-1)^{i+j} \det(M_{j,i})}{\det(A)} b_j$$

but this is exactly $(1/\det(A))$ times the formula for expanding the determinant matrix obtained from A by replacing the i th column with \mathbf{b} . Thus

$$x_i = \frac{\det(\text{matrix obtained from } A \text{ by replacing the } i\text{th column with } \mathbf{b})}{\det(A)}$$

4.7.16 Problems

Problem 4.47 Find the inverse of $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 3 \\ 3 & 0 & 1 \end{bmatrix}$ using the “impractical” formula.

Problem 4.48 Solve the equation

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

using Cramer’s rule.

4.8 Solutions to Chapter Problems

Solution 4.1

$$\begin{aligned} AB &= \begin{bmatrix} -13 & 7 \\ -9 & 5 \end{bmatrix} \\ BA &= \begin{bmatrix} 1 & 2 & -1 \\ -2 & -4 & -8 \\ -1 & -2 & -5 \end{bmatrix} \\ AD &= \begin{bmatrix} -14 \\ -18 \end{bmatrix} \\ CB &= [4 \quad 2] \\ CD &= [26] \\ DC &= \begin{bmatrix} 4 & -4 & 0 \\ -22 & 22 & 0 \\ 4 & -4 & 0 \end{bmatrix} \end{aligned}$$

AC , CA , DA , BC , BD and DB are not defined.

Solution 4.2

$$A = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}; \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; \quad C = \begin{bmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{bmatrix};$$

The matrix A is 3×2 , the matrix B is 2×2 , and the matrix C is 2×3 . For a product of two matrices to be possible, the number of columns in the left matrix has to be equal to the number of rows in the right matrix. Hence there are 5 possible products:

$$\begin{aligned} 1. \quad A \cdot B &= \begin{bmatrix} 0 & 3 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} \\ 2. \quad A \cdot C &= \begin{bmatrix} 3 & 12 & 6 \\ 5 & -2 & 8 \\ 4 & 5 & 7 \end{bmatrix} \\ 3. \quad B \cdot B &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ 4. \quad B \cdot C &= \begin{bmatrix} 3 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix} \\ 5. \quad C \cdot A &= \begin{bmatrix} 1 & 10 \\ 13 & 7 \end{bmatrix} \end{aligned}$$

Solution 4.3 If

$$A = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}$$

then

$$A^2 = \begin{bmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

If

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

then

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

$$A^3 = \begin{bmatrix} 1 & 0 & 3a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Solution 4.4

a,b) $A^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$

c) Use the power series formula

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots$$

and substitute in the matrix tA for each occurrence of x . Substitute the identity matrix for 1. This gives

$$\begin{aligned} e^{tA} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + t \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + \frac{t^2}{2} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} + \dots \\ &= \begin{bmatrix} 1 + t + \frac{t^2}{2} + \frac{t^3}{3!} + \dots & t + 2\frac{t^2}{2} + 3\frac{t^3}{3!} + \dots \\ 0 & 1 + t + \frac{t^2}{2} + \frac{t^3}{3!} + \dots \end{bmatrix} \\ &= \begin{bmatrix} 1 + t + \frac{t^2}{2} + \frac{t^3}{3!} + \dots & t + t^2 + \frac{t^3}{2!} + \dots \\ 0 & 1 + t + \frac{t^2}{2} + \frac{t^3}{3!} + \dots \end{bmatrix} \\ &= \begin{bmatrix} e^t & te^t \\ 0 & e^t \end{bmatrix} \end{aligned}$$

d,e) We are looking for all matrices that satisfy $B^2 = A$. Let $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Then $B^2 = \begin{bmatrix} a^2 + bc & ab + bd \\ ca + dc & bc + d^2 \end{bmatrix}$ so we need to satisfy the equations

$$\begin{aligned} a^2 + bc &= 1 \\ ab + bd &= 1 \\ ca + dc &= 0 \\ bc + d^2 &= 1 \end{aligned}$$

The third equation says $c(a + d) = 0$ so either $c = 0$ or $a + d = 0$. But $a + d = 0$ would contradict the second equations, so we must have $c = 0$. So

$$\begin{aligned} a^2 &= 1 \\ b(a + d) &= 1 \\ d^2 &= 1 \end{aligned}$$

So $a = \pm 1$ and $d = \pm 1$. If $a = 1$ then to satisfy the second equation we must have $d = 1$ and $b = 1/2$. If $a = -1$ then to satisfy the second equation we must have $d = -1$ and $b = -1/2$. Thus the two square roots are $\begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} -1 & -1/2 \\ 0 & -1 \end{bmatrix}$. In general, a matrix may have more than two square roots.

Solution 4.5

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

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

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

$$A^4 = A^5 = \dots = 0$$

Solution 4.6 $T(\mathbf{x} + \mathbf{y}) = \mathbf{x} + \mathbf{y} + \mathbf{a}$ whereas $T(\mathbf{x}) + T(\mathbf{y}) = \mathbf{x} + \mathbf{a} + \mathbf{y} + \mathbf{a} = \mathbf{x} + \mathbf{y} + 2\mathbf{a}$. Since these are not equal, T is not linear.

Solution 4.7 It follows from the properties of the dot product that $\mathbf{a} \cdot (s\mathbf{x} + t\mathbf{y}) = s\mathbf{a} \cdot \mathbf{x} + t\mathbf{a} \cdot \mathbf{y}$.

Solution 4.8 One way to do these problems is to determine the angle θ that the lines make with the x axis, and then substitute into the formula. So, for part (a), we have $\theta = \pi/4$ (45°). Thus the projection matrix is

$$\frac{1}{2} \begin{bmatrix} 1 + \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & 1 - \cos(2\theta) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 + \cos(\pi/2) & \sin(\pi/2) \\ \sin(\pi/2) & 1 - \cos(\pi/2) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Another way to do this problem is to go back to the derivation of the projection matrix, and redo the formula for the matrix for projection in the direction of \mathbf{a} when \mathbf{a} is not necessarily a unit vector. This gives

$$\frac{1}{a_1^2 + a_2^2} \begin{bmatrix} a_1^2 & a_1 a_2 \\ a_2 a_1 & a_2^2 \end{bmatrix}.$$

For part (a) the vector $\mathbf{a} = [1, 1]$ so this formula gives the same answer. For part (b) the vector $\mathbf{a} = [4, -3]$ so the projection matrix is

$$\frac{1}{25} \begin{bmatrix} 16 & -12 \\ -12 & 9 \end{bmatrix}.$$

Solution 4.9 Since we have already computed the corresponding projection matrices, we can use the formula $2P - I$ to get the reflection matrices. This gives

(a)

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

(b)

$$\begin{bmatrix} 7/25 & -24/25 \\ -24/25 & -7/25 \end{bmatrix}$$

Solution 4.10 Here we can just substitute into the formula for the rotation matrix. This gives

(a)

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

(b)

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

(c)

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

Solution 4.11 We need to compute the matrix product

$$\begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} \begin{bmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{bmatrix}$$

This equals

$$\begin{bmatrix} \cos(2\theta)\cos(2\phi) + \sin(2\theta)\sin(2\phi) & \cos(2\theta)\sin(2\phi) - \sin(2\theta)\cos(2\phi) \\ \sin(2\theta)\cos(2\phi) - \cos(2\theta)\sin(2\phi) & \sin(2\theta)\sin(2\phi) + \cos(2\theta)\cos(2\phi) \end{bmatrix}$$

Using the addition formulae for cos and sin this can be rewritten

$$\begin{bmatrix} \cos(2\theta - 2\phi) & -\sin(2\theta - 2\phi) \\ \sin(2\theta - 2\phi) & \cos(2\theta - 2\phi) \end{bmatrix}$$

This is the matrix for rotation by $2\theta - 2\phi$.

Solution 4.12 We first find the angle θ between the line $x = 2y$ and the x-axis.

Note that (using Pythagoras) $\cos \theta = \frac{2}{\sqrt{5}}$, $\sin \theta = \frac{1}{\sqrt{5}}$.

Now,

$$f\left(\begin{bmatrix} 1 \\ 10 \end{bmatrix}\right) = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

So, we need to calculate $\cos(2\theta)$ and $\sin(2\theta)$.

We have $\cos(2\theta) = \cos^2 \theta - \sin^2 \theta = \frac{4}{5} - \frac{1}{5} = \frac{3}{5}$, and $\sin(2\theta) = 2 \sin \theta \cos \theta = 2 \frac{2}{\sqrt{5}} \frac{1}{\sqrt{5}} = \frac{4}{5}$.

So,

$$f\left(\begin{bmatrix} 1 \\ 10 \end{bmatrix}\right) = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix} \begin{bmatrix} 1 \\ 10 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 10 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 43 \\ -26 \end{bmatrix}.$$

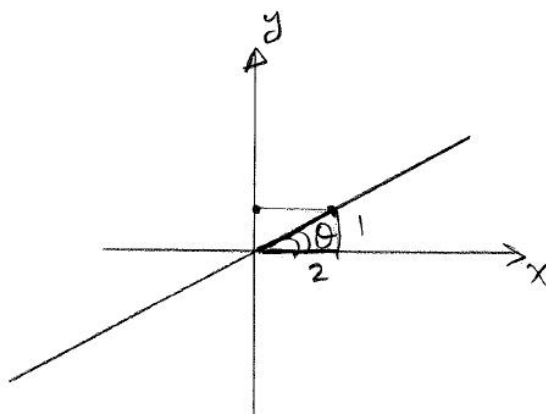


Figure 4.12: Problem 4.12.

Therefore, the matrix of f is

$$\frac{1}{5} \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}.$$

Solution 4.13 Notice that g is in fact the composition of Ref_{θ} and $R_{\pi/2}$, that is:

$$g = R_{\pi/2} \circ Ref_{\theta}$$

where θ is the angle between the line $x = -y$ and the x-axis (so, $\theta = \pi/2 + \pi/4 = 3\pi/4$).

The matrix of g is given by the product of the matrices $R_{\pi/2}$ and $Ref_{3\pi/4}$. Note that

$$\begin{aligned} [g] &= [R_{\pi/2}][Ref_{3\pi/4}] \\ &= \begin{bmatrix} \cos \pi/2 & -\sin \pi/2 \\ \sin \pi/2 & \cos \pi/2 \end{bmatrix} \begin{bmatrix} \cos 6\pi/4 & \sin 6\pi/4 \\ \sin 6\pi/4 & -\cos 6\pi/4 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{aligned}$$

Solution 4.14 Under this rotation, the vector $\mathbf{e}_1 = [1, 0, 0]$ gets transformed to $[\cos(\theta), \sin(\theta), 0]$, $\mathbf{e}_2 = [0, 1, 0]$ gets transformed to $[-\sin(\theta), \cos(\theta), 0]$, and $\mathbf{e}_3 = [0, 0, 1]$ is transformed to itself (em i.e. it doesn't change). Putting this in the columns of a matrix yields

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Solution 4.15 Let

$$\mathbf{x} = (x_1, x_2, x_3, x_4)$$

$$\mathbf{y} = (y_1, y_2, y_3, y_4)$$

1.

$$\begin{aligned} T(\mathbf{x} + \mathbf{y}) &= T(x_1 + y_1, x_2 + y_2, x_3 + y_3, x_4 + y_4) \\ &= ((x_1 + y_1) + 4(x_2 + y_2) + 5(x_3 + y_3), 3(x_1 + y_1) - 2(x_2 + y_2) + \dots \\ &\quad (x_3 + y_3) - (x_4 + y_4), -(x_1 + y_1) - (x_3 + y_3) + (x_4 + y_4)) \\ &= (x_1 + 4x_2 + 5x_3 + y_1 + 4y_2 + 5y_3, 3x_1 - 2x_2 + x_3 - x_4 + \dots \\ &\quad 3y_1 - 2y_2 + y_3 - y_4, -x_1 - x_3 + x_4 - y_1 - y_3 + y_4) \\ &= (x_1 + 4x_2 + 5x_3, 3x_1 - 2x_2 + x_3 - x_4, -x_1 - x_3 + x_4) + \dots \\ &\quad (y_1 + 4y_2 + 5y_3, 3y_1 - 2y_2 + y_3 - y_4, -y_1 - y_3 + y_4) \\ &= T(\mathbf{x}) + T(\mathbf{y}). \end{aligned}$$

This holds for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^4$

2.

$$\begin{aligned} T(c\mathbf{x}) &= T(cx_1, cx_2, cx_3, cx_4) \\ &= (cx_1 + 4cx_2 + 5cx_3, 3cx_1 - 2cx_2 + cx_3 - cx_4, -cx_1 - cx_3 + cx_4) \\ &= (c(x_1 + 4x_2 + 5x_3), c(3x_1 - 2x_2 + x_3 - x_4), c(-x_1 - x_3 + x_4)) \\ &= c(x_1 + 4x_2 + 5x_3, 3x_1 - 2x_2 + x_3 - x_4, -x_1 - x_3 + x_4) \\ &= cT(\mathbf{x}). \end{aligned}$$

This holds for any $\mathbf{x} \in \mathbb{R}^4$, and any $c \in \mathbb{R}$. It is clear then that T is linear.

Solution 4.16 First recall the definition: a vector $\mathbf{y} = (y_1, y_2, y_3)$ is said to be in the range of T if there exists a vector $\mathbf{x} = (x_1, x_2)$ such that $T(\mathbf{x}) = \mathbf{y}$.

a) We should be looking for $\mathbf{x} = (x_1, x_2)$, if any, such that $T(\mathbf{x}) = (1, 4, 2)$. But

$$\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} = T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ x_2 \\ x_1 + x_2 \end{bmatrix}.$$

We get the following system of equations:

$$\begin{cases} x_1 + 2x_2 = 1 \\ x_2 = 4 \\ x_1 + x_2 = 2 \end{cases} \Rightarrow \begin{cases} x_1 + 8 = 1 \\ x_1 + 4 = 2 \end{cases} \Rightarrow \begin{cases} x_1 = -7 \\ x_1 = -2 \end{cases}$$

So, the system does not have any solutions.

In other words, there is no $\mathbf{x} = (x_1, x_2)$ that satisfy the system above. So there is no \mathbf{x} such that $T(\mathbf{x}) = (1, 4, 2)$. Hence, $(1, 4, 2)$ is not in the range of T .

b) We are looking for $\mathbf{x} = (x_1, x_2)$, if any, such that $T(\mathbf{x}) = (1, 1, 1)$. So

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ x_2 \\ x_1 + x_2 \end{bmatrix}.$$

We get the following system of equations:

$$\begin{cases} x_1 + 2x_2 = 1 \\ x_2 = 1 \\ x_1 + x_2 = 1 \end{cases} \Rightarrow \begin{cases} x_1 + 2 = 1 \\ x_1 + 1 = 2 \end{cases} \Rightarrow \begin{cases} x_1 = -1 \\ x_1 = 0 \end{cases}$$

So, as in the previous case, the system does not have any solutions; in other words, $(1, 1, 1)$ is not in the range of T .

Solution 4.17

a) Note that

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

So,

$$\begin{aligned} T\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right) &= 3T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) - 2T\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\right) + 3T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) \\ &= 3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} -1 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} -4 \\ 18 \\ 18 \end{bmatrix} \end{aligned}$$

b) In order to write the matrix of T , we need to find $T(0, 1, 0)$. Note that

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

So,

$$\begin{aligned} T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) &= T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) - T\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}\right) \\ &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}. \end{aligned}$$

Now, the matrix of T is given by

$$T = \left[T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) \quad T\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) \quad T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) \right] = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 5 \\ 1 & 1 & 5 \end{bmatrix}$$

Solution 4.18

a)

$$(h \circ g) \begin{bmatrix} 1 \\ 0 \\ -5 \end{bmatrix} = h \left(g \begin{bmatrix} 1 \\ 0 \\ -5 \end{bmatrix} \right) = h \left(\begin{bmatrix} 1+5+0 \\ -5-2 \times 0 \end{bmatrix} \right) \quad (4.17)$$

$$= h \begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ -5 \\ 5 \end{bmatrix}. \quad (4.18)$$

b) We first find the matrix of g . To do so, we need to know $g(1, 0, 0)$, $g(0, 1, 0)$, and $g(0, 0, 1)$.

We have that

$$g \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad g \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad g \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad .$$

So, the matrix of g is

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

Now we find the matrix of h . As with g , we need to find $h(1, 0)$, and $h(0, 1)$.

We have that

$$h \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad h \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

So, the matrix of h is

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

c) The matrix of $h \circ g$ is given by

$$h \circ g = [h]_{4 \times 2} [g]_{2 \times 3} = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 0 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ -1 & -1 & 1 \\ 0 & -2 & 1 \\ 0 & 2 & -1 \end{bmatrix}_{4 \times 3}$$

We can confirm them that $h \circ g : \mathbb{R}^3 \rightarrow \mathbb{R}^4$. So, the matrix of $h \circ g$ will have 4 rows and 3 columns, as expected.

Solution 4.19 We can determine the transition matrix, since we know each $p_{i,i} = 0$, and the sum over each column is one. We are given that $p_{2,1} = 1/2$. Then, since $p_{1,1} + p_{2,1} + p_{3,1} = 1$, we have $0 + 1/2 + p_{3,1} = 1$, so $p_{3,1} = 1/2$. Similarly $p_{1,2} = 1/3$, $p_{3,2} = 2/3$, and $p_{1,3} = 1/4$, $p_{2,3} = 3/4$. Thus

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

If we start out with the probability vector $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ (i.e., the walker is in location 1) then after two time steps the probability vector is $P^2\mathbf{x}_0$, i.e.,

$$\mathbf{x}_2 = \begin{bmatrix} 7/24 & 1/6 & 1/4 \\ 3/8 & 2/3 & 1/8 \\ 1/3 & 1/6 & 5/8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 7/24 \\ 3/8 \\ 1/3 \end{bmatrix}$$

So the probability that the walker is in position 2 after two time steps is $3/8$.

Solution 4.20

a) Since $P_{i,i} = 0$, and the sum of entries in each column is 1, the matrix P is

$$P = \begin{bmatrix} 0 & 2/5 & 1/4 \\ 1/3 & 0 & 3/4 \\ 2/3 & 3/5 & 0 \end{bmatrix}.$$

b) The random walker starts at location 1. So $\mathbf{x}_0 = (1, 0, 0)^T$. The positions after three time steps will be

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} 0 \\ 1/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} 3/10 \\ 1/2 \\ 1/5 \end{bmatrix}, \quad \mathbf{x}_3 = P\mathbf{x}_2 = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/2 \end{bmatrix},$$

So, $\mathbf{x}_{3,2}$, that is, the probability that the random walker is in location 2 after 3 steps, will be $\frac{1}{4} = 0.25$.

c)

$$\mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} 1/4 \\ 3/4 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = P\mathbf{x}_1 = \frac{1}{60} \begin{bmatrix} 18 \\ 5 \\ 37 \end{bmatrix},$$

So, $\mathbf{x}_{2,1}$, the probability that the random walker is in location 1 after 2 steps, will be $\frac{18}{60} = 0.333\dots$

Solution 4.21

a)

$$P \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (4.19)$$

$$P^2 = \frac{1}{4} \frac{1}{4} \begin{bmatrix} 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & 4 \end{bmatrix} = \frac{1}{4 \cdot 4} \cdot 4 \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}. \quad (4.20)$$

So, $P^2 = P$, thus $P^3 = P \cdot P^2 = P \cdot P = P$. Hence, $P^n = P$ for every $n \geq 1$

b) Let

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Then

$$\mathbf{x}_1 = \frac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Hence, $\mathbf{x}_n = P^n \mathbf{x}_0$. But $P^n = P$, for every $n \geq 1$. So,

$$\mathbf{x}_n = P^n \mathbf{x}_0 = P \mathbf{x}_0 = \mathbf{x}_1 = \frac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Now, let

$$\mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_1 = P \mathbf{x}_0 = \frac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Again

$$\mathbf{x}_n = P^n \mathbf{x}_0 = P \mathbf{x}_0 = \frac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Solution 4.22

$$P = \begin{bmatrix} 1/3 & 1/4 & 2/5 \\ 0 & 1/2 & 1/5 \\ 2/3 & 1/4 & 2/5 \end{bmatrix}$$

a) We are given

$$x_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Then

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} 1/4 \\ 1/2 \\ 1/4 \end{bmatrix}.$$

So, $x_{1,1} = \frac{1}{4}$ is the probability that the random walker is location 1 after one time step.

b) Again

$$x_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Then

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} 1/4 \\ 1/2 \\ 1/4 \end{bmatrix}, \quad \mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} \frac{37}{120} \\ \frac{3}{10} \\ \frac{47}{120} \end{bmatrix} \Rightarrow \mathbf{x}_{2,1} = \frac{37}{120}.$$

c)

$$\mathbf{x}_0 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}.$$

We have

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} \frac{59}{180} \\ \frac{7}{30} \\ \frac{79}{180} \end{bmatrix}, \quad \mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} \frac{1853}{5400} \\ \frac{46}{225} \\ \frac{2443}{5400} \end{bmatrix}$$

Therefore, $\mathbf{x}_{2,3} = \frac{2443}{5400}$

Solution 4.23 In this situation the transition matrix is

$$P = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

so $P = P^2 = P^3 = \dots = P^n$. In this case, if we start out with any probability vector $\mathbf{x}_0 = [x_1 \ x_2 \ x_3]^T$ with $x_1 + x_2 + x_3 = 1$, then

$$P\mathbf{x}_0 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = P^n\mathbf{x}_0$$

for each n .

Solution 4.24

(a) The transition matrix is

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

The system starts in state 3 so $\mathbf{x}^{(0)} = \mathbf{e}_3$. After one step, the probability vector is

$$\mathbf{x}^{(1)} = P\mathbf{x}^{(0)} = \begin{bmatrix} 1/2 & 0 & 1/3 \\ 1/4 & 1/2 & 1/3 \\ 1/4 & 1/2 & 1/3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

That is, there is an equal probability of being in any state. After another step,

$$\mathbf{x}^{(2)} = P\mathbf{x}^{(1)} = \begin{bmatrix} 1/2 & 0 & 1/3 \\ 1/4 & 1/2 & 1/3 \\ 1/4 & 1/2 & 1/3 \end{bmatrix} \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 5/18 \\ 13/36 \\ 13/36 \end{bmatrix}$$

So the probability that the system is in state 2 is $13/36$.

(b) Suppose that the initial probabilities are given by the vector

$$\mathbf{x}^{(0)} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Then after k steps, the probabilities will be $\mathbf{x}^{(k)} = P^k\mathbf{x}^{(0)}$. After many steps (using the limiting behaviour of P^k stated in the question):

$$\begin{aligned} \mathbf{x}^{(k)} &\approx \begin{bmatrix} 0.25 & 0.25 & 0.25 \\ 0.375 & 0.375 & 0.375 \\ 0.375 & 0.375 & 0.375 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \begin{bmatrix} 0.25x_1 + 0.25x_2 + 0.25x_3 \\ 0.375x_1 + 0.375x_2 + 0.375x_3 \\ 0.375x_1 + 0.375x_2 + 0.375x_3 \end{bmatrix} \\ &= (x_1 + x_2 + x_3) \begin{bmatrix} 0.25 \\ 0.375 \\ 0.375 \end{bmatrix} \\ &= \begin{bmatrix} 0.25 \\ 0.375 \\ 0.375 \end{bmatrix} \end{aligned}$$

where the last line follows because the initial probabilities must sum to one. Note that the state tends to this probability distribution no matter what the initial state is (or initial probabilities are). This is known as the steady state probability vector of this random walk.

Solution 4.25

$$A\mathbf{x} = \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{bmatrix}$$

so $\mathbf{y} \cdot A\mathbf{x} = y_1x_1 + 2y_1x_2 + 3y_1x_3 + 4y_2x_1 + 5y_2x_2 + 6y_2x_3$. On the other hand

$$A^T\mathbf{y} = \begin{bmatrix} y_1 + 4y_2 \\ 2y_1 + 5y_2 \\ 3y_1 + 6y_2 \end{bmatrix}$$

so $(A^T\mathbf{y}) \cdot \mathbf{x} = y_1x_1 + 4y_2x_1 + 2y_1x_2 + 5y_2x_2 + 3y_1x_3 + 6y_2x_3$. These expressions are equal.

Solution 4.26 $(A^T)^T = A$.

Solution 4.27 $AB = \begin{bmatrix} 9 & 12 & 15 \\ 7 & 11 & 15 \end{bmatrix}$ so $(AB)^T = \begin{bmatrix} 9 & 7 \\ 12 & 11 \\ 15 & 15 \end{bmatrix}$. On the other

hand $B^T A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 7 \\ 12 & 11 \\ 15 & 15 \end{bmatrix}$.

Solution 4.28 If $\mathbf{y} = \mathbf{e}_i$ and $\mathbf{x} = \mathbf{e}_j$ (vectors that are all zeros except in one spot) then $\mathbf{y} \cdot (A\mathbf{x}) = \mathbf{e}_i \cdot (A\mathbf{e}_j)$ is the matrix entry a_{ij} . So we can conclude that all the matrix entries of A are the same as those for B . This means the matrices must be the same.

Solution 4.29 $\mathbf{x} \cdot (A\mathbf{y}) = \mathbf{x}^T A\mathbf{y} = (A^T\mathbf{x})^T \mathbf{y} = (A^T\mathbf{x}) \cdot \mathbf{y}$

Solution 4.30 To determine whether these matrices invertible we reduce them using Gaussian elimination. This gives

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

(b) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 0 & -1/3 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 0 & 2/3 \end{bmatrix}$

Since these all have rank equal to their size, they are all invertible.

Solution 4.31 The inverse is

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

Solution 4.32

(a) This matrix reduces to $\begin{bmatrix} 2 & 3 & -1 \\ 0 & 1/2 & 7/2 \\ 0 & 0 & 0 \end{bmatrix}$ and so is not invertible

(b) The inverse is $\begin{bmatrix} -1 & 0 & 1 \\ 1 & 1 & 0 \\ 3 & 1 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} 3 & -5/2 & 1/2 \\ -3 & 4 & -1 \\ 1 & -3/2 & 1/2 \end{bmatrix}$

(d) $\begin{bmatrix} -7 & 5 & 3 \\ 3 & -2 & -2 \\ 3 & -2 & -1 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 0 & -a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(f) $\begin{bmatrix} 1 & -a & ac - b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{bmatrix}$

Solution 4.33 In order to find the inverse of A , we form the augmented matrix $[A|I]$. Then we apply elementary row operations to A , to reduce A to its rref. If A is invertible, then using this method $[A|I]$ transfers to $[I|B]$, where B is the inverse of A .

Now, let

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{bmatrix}$$

and form the augmented matrix:

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

The reduced row echelon form of the augmented matrix is

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 13/8 & -1/2 & -1/8 \\ 0 & 1 & 0 & -15/8 & 1/2 & 3/8 \\ 0 & 0 & 1 & 5/4 & 0 & -1/4 \end{array} \right]$$

Therefore, the inverse of A is

$$A^{-1} = \begin{bmatrix} 13/8 & -1/2 & -1/8 \\ -15/8 & 1/2 & 3/8 \\ 5/4 & 0 & -1/4 \end{bmatrix}.$$

Let

$$B = \begin{bmatrix} 1 & 2 & -3 & 1 \\ -1 & 3 & -3 & -2 \\ 2 & 0 & 1 & 5 \\ 3 & 1 & -2 & 5 \end{bmatrix}$$

Applying the same method, we observe that B is not invertible because

$$\text{rref}(B) = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note that rank of $B = 3$, which is the number of non-zero rows in $\text{rref}(B)$.

Solution 4.34

a) Let

$$A = \begin{bmatrix} 1 & 0 & 5 \\ 1 & -2 & 3 \\ 2 & 1 & -3 \end{bmatrix}$$

be the coefficient matrix. Then

$$A\mathbf{x} = \begin{bmatrix} 6 \\ 14 \\ -2 \end{bmatrix}$$

b)

$$A^{-1} = \begin{bmatrix} 3/28 & 5/28 & 5/14 \\ 9/28 & -13/28 & 1/14 \\ 5/28 & -1/28 & -1/14 \end{bmatrix} = \frac{1}{28} \begin{bmatrix} 3 & 5 & 10 \\ 9 & -13 & 2 \\ 5 & -1 & -2 \end{bmatrix}$$

c) In order to solve the system $A\mathbf{x} = \mathbf{b}$, we multiply both sides by A^{-1} :

$$\begin{aligned} A^{-1}A\mathbf{x} &= A^{-1}\mathbf{b} = A^{-1} \begin{bmatrix} 6 \\ 14 \\ -2 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} 17 \\ -33 \\ 5 \end{bmatrix} \\ \Rightarrow \quad I\mathbf{x} &= \mathbf{x} = \begin{bmatrix} 17 \\ -33 \\ 5 \end{bmatrix}. \end{aligned}$$

So, the system has the **unique** solution

$$\mathbf{x} = \begin{bmatrix} 17 \\ -33 \\ 5 \end{bmatrix}$$

Solution 4.35 We reduce the matrix as follows:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 7 \\ 0 & 2 & 8 & 26 \\ 0 & 3 & 15 & 63 \end{bmatrix} \begin{array}{l} \\ (R2) - (R1) \\ (R3) - (R1) \\ (R4) - (R1) \end{array}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 7 \\ 0 & 0 & 2 & 12 \\ 0 & 0 & 6 & 42 \end{bmatrix} \begin{array}{l} \\ \\ (R3) - 2(R2) \\ (R4) - 3(R2) \end{array}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 7 \\ 0 & 0 & 2 & 12 \\ 0 & 0 & 0 & 6 \end{bmatrix} \begin{array}{l} \\ \\ \\ (R4) - 3(R3) \end{array}$$

None of these operations affect the determinant. So the determinant of the original matrix is the same as the determinant of the reduced diagonal matrix. This determinant is the product of the diagonal elements which equals 12.

Solution 4.36 We reduce the matrix as follows:

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 2 & 4 & 8 \\ 1 & -2 & 4 & -8 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 3 & 3 & 9 \\ 0 & -1 & 3 & -7 \\ 0 & 2 & 0 & 2 \end{bmatrix} \begin{array}{l} \\ (R2) - (R1) \\ (R3) - (R1) \\ (R4) - (R1) \end{array}$$

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & 1 & 3 \\ 0 & -1 & 3 & -7 \\ 0 & 2 & 0 & 2 \end{bmatrix} \begin{array}{l} \\ (1/3)(R2) \\ \\ \end{array}$$

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 4 & -4 \\ 0 & 0 & -2 & -4 \end{bmatrix} \begin{array}{l} \\ \\ (R3) + (R2) \\ (R4) - 2(R2) \end{array}$$

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -2 & -4 \end{bmatrix} \begin{array}{l} \\ \\ (1/4)(R3) \\ \end{array}$$

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -6 \end{bmatrix} (R4) + 2(R3)$$

The two operations that changed the determinant were multiplying the second row by $1/3$ and multiplying the third row by $1/4$. Thus the determinant of the diagonal matrix is $(1/3)(1/4) \times$ the determinant of the original matrix. Hence the determinant of the original matrix is $3 \times 4 \times (-6) = -72$.

Solution 4.37

a)

$$\begin{aligned} \det A &= \det \begin{bmatrix} 2 & 0 & 2 & 4 \\ 0 & 0 & 3 & 2 \\ 2 & 2 & 4 & 4 \\ 3 & 0 & 6 & 2 \end{bmatrix} \\ &= 2 \det \begin{bmatrix} 0 & 3 & 2 \\ 2 & 4 & 4 \\ 0 & 6 & 2 \end{bmatrix} - 0 + 2 \det \begin{bmatrix} 0 & 0 & 2 \\ 2 & 2 & 4 \\ 3 & 0 & 2 \end{bmatrix} - 4 \det \begin{bmatrix} 0 & 0 & 3 \\ 2 & 2 & 4 \\ 3 & 0 & 6 \end{bmatrix} \\ &= 2 \left(0 - 3 \det \begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix} + 2 \det \begin{bmatrix} 2 & 4 \\ 0 & 6 \end{bmatrix} \right) + 2(2) \det \begin{bmatrix} 2 & 2 \\ 3 & 0 \end{bmatrix} - \dots \\ &\quad 4(3) \det \begin{bmatrix} 2 & 2 \\ 3 & 0 \end{bmatrix} \\ &= 2(0 - 3(4 - 0) + 2((2 - 0)) + 4(0 - 6) - 12(0 - 6)) \\ &= 2(12) - 24 + 72 = \mathbf{72}. \end{aligned}$$

b)

$$\begin{aligned} \det A &= \det \begin{bmatrix} 2 & 0 & 2 & 4 \\ 0 & 0 & 3 & 2 \\ 2 & 2 & 4 & 4 \\ 3 & 0 & 6 & 2 \end{bmatrix} \\ &= -0 + 0 - 2 \det \begin{bmatrix} 2 & 2 & 4 \\ 0 & 3 & 2 \\ 3 & 6 & 2 \end{bmatrix} + 0 \\ &= -2 \left(-0 + 3 \det \begin{bmatrix} 2 & 4 \\ 3 & 2 \end{bmatrix} - 2 \det \begin{bmatrix} 2 & 2 \\ 3 & 6 \end{bmatrix} \right) \\ &= -2(3(4 - 12) - 2(12 - 6)) \\ &= -2(-24 - 12) = \mathbf{72} \end{aligned}$$

Solution 4.38

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

$$\begin{aligned}
 &= -1 \times 0 + 2 \times (1 - 3) - 3 \times 0 = -4 \\
 \det \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 3 \\ 3 & 0 & 1 \end{bmatrix} &= 1 \det \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} - 3 \det \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix} + 1 \det \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \\
 &= 1 \times (0 - 6) - 3 \times 0 + 1 \times (2 - 0) = -4
 \end{aligned}$$

Solution 4.39

$$\begin{aligned}
 D &= \det \begin{bmatrix} 2 & 0 & 2 & 4 \\ 0 & 0 & 3 & 2 \\ 2 & 2 & 4 & 4 \\ 3 & 0 & 6 & 2 \end{bmatrix} &\rightarrow (1, :) = \frac{1}{2}(1, :) \\
 \frac{1}{2}D &= \det \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 0 & 3 & 2 \\ 2 & 2 & 4 & 4 \\ 3 & 0 & 6 & 2 \end{bmatrix} &\rightarrow \begin{aligned} (3, :) &= (3, :) - 2(1, :) \\ (4, :) &= (4, :) - 3(1, :) \end{aligned} \\
 \frac{1}{2}D &= \det \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 0 & 3 & 2 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 3 & -4 \end{bmatrix} &\rightarrow (4, :) = (4, :) - (2, :) \\
 \frac{1}{2}D &= \det \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 0 & 3 & 2 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & -6 \end{bmatrix} &\rightarrow (2, :) \leftrightarrow (3, :) \\
 -\frac{1}{2}D &= \det \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & -6 \end{bmatrix}
 \end{aligned}$$

Therefore, $-\frac{1}{2}D = (1)(2)(3)(-6) = -36 \rightarrow D = \mathbf{72}$.

Solution 4.40 If n is even, then $\frac{n}{2}$ row interchanges are required to reduce the matrix to a diagonal matrix. Therefore, the determinant is $(-1)^{n/2} a_1 a_2 \cdots a_n$.

If n is odd, then $\frac{n-1}{2}$ row interchanges are required to reduce the matrix to a diagonal matrix (note that the centre row stays). Therefore, the determinant is $(-1)^{\frac{n-1}{2}} a_1 a_2 \cdots a_n$.

In other words:

$$\det \begin{bmatrix} & & & a_n \\ & & \ddots & \\ & & & \\ a_2 & & & \\ a_1 & & & \end{bmatrix} = (-1)^k \det \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix} = (-1)^k a_1 a_2 \cdots a_n,$$

where

$$k = \begin{cases} \frac{n}{2} & n \text{ even} \\ \frac{n-1}{2} & n \text{ odd} \end{cases}$$

Alternatively, we can expand the determinant on the first row:

$$\begin{aligned}
 \det \begin{bmatrix} & & & a_n \\ & & \ddots & \\ & a_2 & & \\ a_1 & & & \end{bmatrix} &= (-1)^{1+n} a_n \det \begin{bmatrix} & & & a_{n-1} \\ & & \ddots & \\ & a_2 & & \\ a_1 & & & \end{bmatrix} \\
 &= (-1)^{1+n} a_n (-1)^{1+(n-1)} a_{n-1} \det \begin{bmatrix} & & & a_{n-2} \\ & & \ddots & \\ & a_2 & & \\ a_1 & & & \end{bmatrix} \\
 &= \dots \\
 &= (-1)^{1+n} a_n (-1)^{1+(n-1)} a_{n-1} \dots (-1)^{1+1} a_1 \\
 &= (-1)^{(1+n)+n+\dots+2} a_1 a_2 \dots a_n \\
 &= (-1)^{\frac{(n+1)(n+2)}{2}-1} a_1 a_2 \dots a_n
 \end{aligned}$$

Solution 4.41 A matrix is not invertible if and only if its determinant is equal to zero. Therefore, all we need to do is find the determinant of the matrix, and determine the values of λ that make the determinant zero.

$$\begin{aligned}
 \det \begin{bmatrix} 2-\lambda & 1 & 0 \\ -1 & -\lambda & 1 \\ 1 & 3 & 1-\lambda \end{bmatrix} &= (2-\lambda) \det \begin{bmatrix} -\lambda & 1 \\ 3 & 1-\lambda \end{bmatrix} - \det \begin{bmatrix} -1 & 1 \\ 1 & 1-\lambda \end{bmatrix} + 0 \\
 &= (2-\lambda)(-\lambda(1-\lambda) - 3) - (-(1-\lambda) - 1) \\
 &= (2-\lambda)(\lambda^2 - \lambda - 3) - (\lambda - 2) \\
 &= (2-\lambda)(\lambda^2 - \lambda - 2) \\
 &= (2-\lambda)(\lambda - 2)(\lambda + 1) = 0 \\
 \Rightarrow \quad \lambda &= \mathbf{2, -1}
 \end{aligned}$$

Solution 4.42

a) **False** Counterexample:

$$\det \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = 0 \neq 1$$

The correct statement should be:

The determinant of a **triangular (or diagonal) matrix** is the product of the entries of its main diagonal.

b) **False** Counterexample:

$$A = [1 \ 2] \in \mathbb{R}^{1 \times 2}.$$

Then

$$\begin{aligned}\det(AA^T) &= \det\left([1 \ 2] \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) \\ &= \det([5]) = 5 \\ \det(A^T A) &= \det\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix} [1 \ 2]\right) \\ &= \det\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = 0\end{aligned}$$

Therefore, $\det(AA^T) \neq \det(A^T A)$.

The correct statement should be:

For **every square matrix** A , we have $\det(AA^T) = \det(A^T A)$.

c) **True** Since

$$\begin{aligned}\det(B^{-1}AB) &= \det(B^{-1}) \det(A) \det(B) = \frac{\det(A) \det(B)}{\det(B)} = \det(A) \\ \det(BAB^{-1}) &= \det(B) \det(A) \det(B^{-1}) = \frac{\det(B) \det(A)}{\det(B)} = \det(A).\end{aligned}$$

d) **False** Counterexample: Let

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

Then

$$\begin{aligned}\det(A) &= 1 \\ \det(2A) &= \det\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 4 \\ \Rightarrow \det(A) &\neq 2 \det(A).\end{aligned}$$

The correct statement should be:

If an $n \times n$ matrix A is multiplied by a scalar c , the determinant of the resulting matrix is $\det(c\mathbf{A}) = c^n \cdot \det(\mathbf{A})$.

e) **True** Since suppose the column vectors of $A \in \mathbb{R}^{n \times n}$ are linearly independent, then $A\mathbf{x} = \mathbf{0}$ has a unique solution $\mathbf{x} = \vec{\mathbf{0}}$. The row vectors of A^T will also be linearly independent, hence $A^T\mathbf{x} = \mathbf{0}$ also has a unique solution $\mathbf{x} = \vec{\mathbf{0}}$. Furthermore, since the columns are linearly independent,

$$\det(A) \neq 0 \quad \Rightarrow \quad \det(A^T) = \det(A) \neq 0.$$

Solution 4.43 $B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ so $B^T B = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$ and $B^T \mathbf{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. The equation $B^T B \mathbf{x} = B^T \mathbf{b}$ has solution $\mathbf{x} = \begin{bmatrix} 1 \\ -1/2 \end{bmatrix}$. For this vector \mathbf{x} we have $B\mathbf{x} = \begin{bmatrix} 1/2 \\ 1 \\ 1/2 \end{bmatrix}$.

Solution 4.44 We obtain $B = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ so that $B^T B = \begin{bmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{bmatrix}$ (where $n = 3$ in this example) and $B^T \mathbf{b} = \begin{bmatrix} \sum x_i y_i \\ \sum y_i \end{bmatrix}$. Thus we end up with the same equations as before.

Solution 4.45 The inverse operation to multiplying the first row by two is multiplying the first row by $1/2$. Therefore the inverse elementary matrix to $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is $\begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

The inverse operation to subtracting twice the first row from the second row is adding twice the first row to the second row. Therefore the inverse elementary matrix to $\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

The inverse operation to exchanging the last two rows is exchanging them again. Therefore the inverse elementary matrix to $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ is the same

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

In each case one can check directly that the inverse matrices when multiplied by the original matrices give the identity matrix.

Solution 4.46 We can reduce A to the identity with the following row operations: $(R2) - (1/2)(R1)$, $(R3) + (1/2)(R1)$, $(R3) - (R2)$, $(R1) - 6(R2)$, $(R2) - (5/4)(R3)$, $(R1) + 7(R3)$, $(1/2)(R1)$, $2(R2)$, $(1/2)(R3)$. So

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -5/4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -6 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = I$$

so

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 6 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 5/4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/22 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Solution 4.47 The determinant is -4 by problem 4.38, and

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

Thus the inverse is

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

Solution 4.48 According to Cramer's rule

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

Chapter 5

Complex numbers

5.1 Complex arithmetic

Complex numbers can be thought of as points on the xy plane. The point (x, y) , thought of as a complex number, is written $x + iy$ (or $x + jy$ if you are an electrical engineer). The i stands for an “imaginary” quantity such that $i^2 = -1$. If $z = x + iy$ then x is called the real part of z and y is called the imaginary part of z .

Complex numbers are added just as if they were vectors in two dimensions. If $z = x + iy$ and $w = s + it$, then

$$z + w = (x + iy) + (s + it) = (x + s) + i(y + t)$$

To multiply two complex numbers, just remember that $i^2 = -1$. So if $z = x + iy$ and $w = s + it$, then

$$zw = (x + iy)(s + it) = xs + i^2 ytr + iys + ixt = (xs - yt) + i(xt + ys)$$

The modulus of a complex number, denoted $|z|$ is simply the length of the corresponding vector in two dimensions. If $z = x + iy$

$$|z| = |x + iy| = \sqrt{x^2 + y^2}$$

An important property is

$$|zw| = |z||w|$$

just like for real numbers.

The complex conjugate of a complex number z , denoted \bar{z} , is the reflection of z across the x axis (also called the real axis). Thus $\overline{x + iy} = x - iy$. Thus complex conjugate is obtained by changing all the i 's to $-i$'s. We have

$$\overline{\bar{z}} = z$$

and

$$z\bar{z} = |z|^2$$

This last equality is useful for simplifying fractions of complex numbers by turning the denominator into a real number, since

$$\frac{z}{w} = \frac{z\bar{w}}{|w|^2}$$

For example, to simplify $(1+i)/(1-i)$ we can write

$$\frac{1+i}{1-i} = \frac{(1+i)^2}{(1-i)(1+i)} = \frac{1-1+2i}{2} = i$$

A complex number z is real (i.e. the y part in $x+iy$ is zero) whenever $\bar{z} = z$. We also have the following formulas for the real and imaginary part. If $z = x + iy$ then $x = (z + \bar{z})/2$ and $y = (z - \bar{z})/(2i)$

Complex numbers are indispensable in many practical calculations. We will discuss complex exponentials when we talk about differential equations. The reason why we are interested in them now is the following fact:

Theorem 5.1 *If we use complex numbers, every polynomial can be completely factored.* In other words, given a polynomial $\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$, there exist (possibly complex) numbers r_1, r_2, \dots, r_n such that

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = (\lambda - r_1)(\lambda - r_2) \cdots (\lambda - r_n)$$

The numbers r_1 are the values of λ for which the polynomial is zero.

So for example the polynomial $\lambda^2 + 1$ has no real roots, since there is no real number λ for which it is zero. However there are two complex roots, $\pm i$ and

$$\lambda^2 + 1 = (\lambda + i)(\lambda - i)$$

Of course, actually finding the roots of a high degree polynomial is difficult. Here are some points to keep in mind.

You can always find the roots of a quadratic polynomial using the quadratic formula. In other words, the roots of $a\lambda^2 + b\lambda + c$ are

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

If the quantity inside the square root is negative, then the roots are complex. So, for example the roots of $\lambda^2 + \lambda + 1$ are

$$\frac{-1 \pm \sqrt{1^2 - 4}}{2} = \frac{-1 \pm \sqrt{-3}}{2} = \frac{-1 \pm \sqrt{-1}\sqrt{3}}{2} = \frac{-1}{2} \pm i \frac{\sqrt{3}}{2}.$$

5.2 Complex matrices and linear systems

Determinants can be computed using basic arithmetic operations. Also, linear systems can be solved and matrix inverses found by applying basic arithmetic operations in Gaussian elimination. Having learned the arithmetic of complex numbers, we can apply it to finding determinants and inverses of matrices with complex entries (complex matrices) and solutions of complex linear systems.

5.2.1 Determinants of complex matrices

The determinant of a square matrix with complex entries is calculated in the same way as for a real matrix. As in the real case, a complex matrix is invertible if and only if its determinant is nonzero.

Example 5.1 Calculate the determinant of

$$\mathbf{A} = \begin{bmatrix} 1+i & 2 \\ 3 & 1-2i \end{bmatrix}$$

Using the same formula as for real matrices we compute

$$\det \mathbf{A} = (1+i)(1-2i) - 6 = -3 - i.$$

Since the determinant is not zero, \mathbf{A} is invertible. We will find the unique solution of a system with this coefficient matrix in Example 5.5 and its inverse in Example 5.6 below.

Example 5.2 Calculate the determinant of

$$\mathbf{A} = \begin{bmatrix} i & 0 & 3 \\ 1 & i & -1 \\ 0 & 1 & 3+i \end{bmatrix}$$

We compute

$$\det \mathbf{A} = i[i(3+i) + 1] + 3 = -3 + 3 = 0.$$

In this example, \mathbf{A} is not invertible and so has nontrivial homogeneous solutions. We will find these homogeneous solutions in Example 5.4 below.

5.2.2 Homogeneous complex linear systems

Homogeneous linear systems with complex coefficients have solutions that can be found after Gaussian elimination as was done for real systems in Chapter 3.

Example 5.3 Find the complex numbers x_1 , x_2 and x_3 that satisfy

$$(1+i)x_1 + 2x_2 - 3ix_3 = 0 \quad (5.1)$$

$$(1-i)x_1 + 3x_2 + (1+i)x_3 = 0 \quad (5.2)$$

Note that both right hand side values above are zero: this defines a homogeneous system. Remember that there is always at least one solution with all variables equal to zero. In this example, there will be an infinite number of solutions. We can write this system in an augmented matrix:

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

However, just like real homogeneous systems, we can just remember the zero right hand sides (which will stay zero during the elimination process) and write

$$\begin{aligned} \begin{bmatrix} 1+i & 2 & -3i \\ 1-i & 3 & 1+i \end{bmatrix} &\sim \begin{bmatrix} 1 & 1-i & -\frac{3}{2}-\frac{3}{2}i \\ 0 & 3+2i & 4+i \end{bmatrix} \begin{array}{l} (1) \div (1+i) \\ (2) - (1-i)(\text{new1}) \end{array} \\ &\sim \begin{bmatrix} 1 & 0 & -\frac{57}{26}-\frac{1}{26}i \\ 0 & 1 & \frac{14}{13}-\frac{5}{13}i \end{bmatrix} \begin{array}{l} (1) - (1-i)(\text{new2}) \\ (2) \div (3+2i) \end{array} \end{aligned} \quad (5.3)$$

The final line (5.3) above is the homogeneous system in reduced row echelon form. It can be seen that x_3 is not determined. It is a free variable which we will denote by t , a complex parameter. From the second row of (5.3)

$$x_2 = -\left(\frac{14}{13} - \frac{5}{13}i\right)t$$

and from the first row of (5.3)

$$x_1 = -\left(-\frac{57}{26} - \frac{1}{26}i\right)t$$

or summarizing

$$\mathbf{x} = \left(\frac{57}{26} + \frac{1}{26}i, -\frac{14}{13} + \frac{5}{13}i, 1\right)t$$

is a parametric form of all of the solutions to the system with t any complex number.

Example 5.4 Find a nontrivial solution to the homogeneous system with coefficient matrix

$$\mathbf{A} = \begin{bmatrix} i & 0 & 3 \\ 1 & i & -1 \\ 0 & 1 & 3+i \end{bmatrix}$$

from Example 5.2. We do Gaussian elimination on the matrix:

$$\begin{aligned} \begin{bmatrix} i & 0 & 3 \\ 1 & i & -1 \\ 0 & 1 & 3+i \end{bmatrix} &\sim \begin{bmatrix} 1 & 0 & -3i \\ 0 & i & -1+3i \\ 0 & 1 & 3+i \end{bmatrix} \begin{array}{l} (1) \div i \\ (2) - (\text{new1}) \end{array} \\ &\sim \begin{bmatrix} 1 & 0 & -3i \\ 0 & 1 & 3+i \\ 0 & 0 & 0 \end{bmatrix} \begin{array}{l} (2) \div i \\ (3) - (\text{new2}) \end{array} \end{aligned}$$

All solutions can be written in parametric form

$$\mathbf{x} = (3i, -3 - i, 1)t.$$

We can take $t = 1$ or any $t \neq 0$ to get a nontrivial solution to the homogeneous system.

Note that in Chapter 6, when we are finding eigenvectors corresponding to complex eigenvalues, this is equivalent to finding nontrivial solutions to homogeneous systems with square, complex coefficient matrices with determinant zero just as we did in the example above.

5.2.3 Non-homogeneous complex linear systems

These are solved with Gaussian elimination just as in the real case.

Example 5.5 Find all solutions x_1 and x_2 of the linear system

$$\begin{aligned}(1+i)x_1 + 2x_2 &= 1 \\ 3x_1 + (1-2i)x_2 &= i\end{aligned}$$

Note that the coefficient matrix of the system is the same as in Example 5.1. It has a nonzero determinant so the system above has a unique solution. We write the system in an augmented matrix and perform elimination:

$$\begin{aligned}\left[\begin{array}{cc|c} 1+i & 2 & 1 \\ 3 & 1-2i & i \end{array} \right] &\sim \left[\begin{array}{cc|c} 1 & 1-i & \frac{1}{2} - \frac{1}{2}i \\ 0 & -2+i & -\frac{3}{2} + \frac{5}{2}i \end{array} \right] & \begin{array}{l} (1) \div (1+i) \\ (2) - 3(\text{new1}) \end{array} \\ &\sim \left[\begin{array}{cc|c} 1 & 0 & \frac{1}{10} + \frac{13}{10}i \\ 0 & 1 & \frac{11}{10} - \frac{7}{10}i \end{array} \right] & \begin{array}{l} (1) - (1-i)(\text{new2}) \\ (2) \div (-2+i) \end{array}\end{aligned}$$

The solution $x_1 = 1/10 + 13/10i$, $x_2 = 11/10 - 7/10i$ can be read off the reduced row echelon form above.

5.2.4 Finding inverses of complex matrices

This is also done with Gaussian elimination as in the real case.

Example 5.6 Find the inverse of the matrix in Example 5.1:

$$\mathbf{A} = \begin{bmatrix} 1+i & 2 \\ 3 & 1-2i \end{bmatrix}$$

Proceed as in the real case:

$$\begin{aligned}[\mathbf{A}|\mathbf{I}] &= \left[\begin{array}{cc|cc} 1+i & 2 & 1 & 0 \\ 3 & 1-2i & 0 & 1 \end{array} \right] \\ &\sim \left[\begin{array}{cc|cc} 1 & 1-i & \frac{1}{2} - \frac{1}{2}i & 0 \\ 0 & -2+i & -\frac{3}{2} + \frac{3}{2}i & 1 \end{array} \right] \\ &\sim \left[\begin{array}{cc|cc} 1 & 0 & -\frac{1}{10} + \frac{7}{10}i & \frac{3}{5} - \frac{1}{5}i \\ 0 & 1 & \frac{9}{10} - \frac{3}{10}i & -\frac{2}{5} - \frac{1}{5}i \end{array} \right]\end{aligned}$$

So

$$\mathbf{A}^{-1} = \begin{bmatrix} -\frac{1}{10} + \frac{7}{10}i & \frac{3}{5} - \frac{1}{5}i \\ \frac{9}{10} - \frac{3}{10}i & -\frac{2}{5} - \frac{1}{5}i \end{bmatrix}$$

5.3 Complex exponential

We begin by considering the differential equation

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

In other words we are looking for a function whose derivative is equal to the function. The exponential is such a function, so $y(t) = e^t$ is a solution to this differential equation. So is $y(t) = Ce^t$, where C is a constant.

Now consider the equation

$$y'(t) = ay(t)$$

where a is a real number. Then, using the chain rule, we see that $y(t) = Ce^{at}$ is a solution for any choice of constant C . Notice that the constant C is $y(0)$, the value of the solution at time zero. If we insist that the solution at time $t = 0$ take on a particular value

$$y(0) = y_0$$

Then this forces the constant C to be y_0

Now consider the equation

$$y'(t) = iy(t)$$

A solution to this equation is given by

$$y(t) = \cos(t) + i \sin(t)$$

To check this, just differentiate.

$$\begin{aligned} y'(t) &= -\sin(t) + i \cos(t) \\ &= i(\cos(t) + i \sin(t)) \\ &= iy(t) \end{aligned}$$

So it is natural to define the exponential, e^{it} , of a purely imaginary number it to be

$$e^{it} = \cos(t) + i \sin(t)$$

The complex exponential satisfies the familiar rule $e^{i(s+t)} = e^{is}e^{it}$ since by the addition formulas for sine and cosine

$$\begin{aligned} e^{i(s+t)} &= \cos(s+t) + i \sin(s+t) \\ &= \cos(s)\cos(t) - \sin(s)\sin(t) + i(\sin(s)\cos(t) + \cos(s)\sin(t)) \\ &= (\cos(s) + i \sin(s))(\cos(t) + i \sin(t)) \\ &= e^{is}e^{it} \end{aligned}$$

Now it is easy to check that solutions to

$$y'(t) = iby(t)$$

are given by $y(t) = Ce^{ibt}$, where C is an arbitrary constant. Since we are dealing with complex numbers, we allow C to be complex too.

The exponential of a number that has both a real and imaginary part is defined in the natural way.

$$e^{a+ib} = e^a e^{ib} = e^a (\cos(b) + i \sin(b))$$

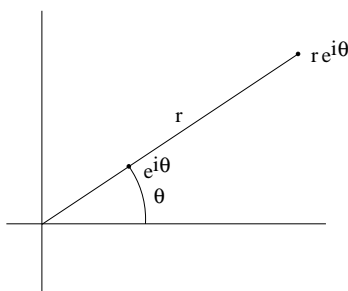


Figure 5.1: Polar representation of a complex number.

and it is easy to check that the solution to the differential equation

$$y'(t) = \lambda y(t) = (a + ib)y(t)$$

is given by $y(t) = Ce^{\lambda t} = Ce^{(a+ib)t}$. As before, if we insist that the solution at time $t = 0$ take on a particular value

$$y(0) = y_0,$$

then this forces the constant C to be y_0 .

5.4 Polar representation of a complex number

Notice that the number $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ lies on the unit circle on the complex plane, at the point making an angle of θ (radians) with the x axis. If we multiply $e^{i\theta}$ by a real number r , then we obtain the complex number whose polar co-ordinates are r and θ . This is shown in Figure 5.1.

Notice that r is exactly the length (sometimes called the modulus) of the complex number $re^{i\theta}$. The angle θ is called the argument. This representation of complex numbers makes the definition of complex multiplication more transparent. We have

$$r_1 e^{i\theta_1} r_2 e^{i\theta_2} = r_1 r_2 e^{i(\theta_1 + \theta_2)}$$

In other words, when we multiply two complex numbers, the moduli get multiplied and the arguments get added.

5.5 MATLAB

Complex numbers are handled naturally by MATLAB. Specifying a complex number is done with the syntax `z = 1+3i`. The commands `real(z)` and `imag(z)` return the real and imaginary parts of a complex number z . The commands `abs(z)` and `conj(z)` return the length and conjugate of z . Typing `sqrt(-1)` returns

ans =

$$0 + 1.0000i$$

MATLAB commands `det`, `rref`, and `inv` all work on complex matrices.

5.6 Problems

Problem 5.1 Show that $|zw| = |z||w|$ for complex numbers z and w .

Problem 5.2 Show that $\overline{z\bar{w}} = \bar{z}w$ for complex numbers z and w .

Problem 5.3 Show that $z\bar{z} = |z|^2$ for every complex numbers z .

Problem 5.4 Simplify the following expressions to the form $x + iy$.

1. $i(2 - 3i)(-2 + i)$

2. $\frac{5}{(1 - i)(2 - i)(3 - i)}$

3. $(-1 + i)^{50}$

Problem 5.5 Prove that

1. $\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|}$;

2. $\overline{(z^n)} = (\bar{z})^n$ for all $n \in \mathbb{N}$;

3. $\frac{1}{z} = \bar{z}$ if $|z| = 1$;

4. z is either real or pure imaginary if $(\bar{z})^2 = z^2$.

Problem 5.6 There are three values of z which gives $z^3 = -i$. Find all three values in the form $x + iy$.

5.7 Solutions to Chapter Problems

Solution 5.1 If $z = x + iy$ and $w = s + it$ then $zw = xs - yt + i(xt + ys)$ so $|zw|^2 = (xs - yt)^2 + (xt + ys)^2 = x^2s^2 + y^2t^2 - 2xyst + x^2t^2 + y^2s^2 + 2xyst = (x^2 + y^2)(s^2 + t^2) = |z|^2|w|^2$.

Solution 5.2 If $z = x + iy$ and $w = s + it$ then $zw = xs - yt + i(xt + ys)$ so $\overline{z\bar{w}} = xs - yt - i(xt + ys)$. On the other hand $\bar{z}w = (x - iy)(s + it) = xs - yt - i(xt + ys)$.

Solution 5.3 If $z = x + iy$ then $z\bar{z} = (x + iy)(x - iy) = x^2 + y^2 = |z|^2$.

Solution 5.4

a)

$$\begin{aligned}
 i(2-3i)(-2+i) &= (3+2i)(-2+i) \\
 &= -6-4i+3i+2i^2 \\
 &= -8-i
 \end{aligned}$$

b)

$$\begin{aligned}
 \frac{5}{(1-i)(2-i)(3-i)} &= \frac{5(1+i)(2+i)(3+i)}{(1-i)(2-i)(3-i)(1+i)(2+i)(3+i)} \\
 &= \frac{5(1+3i)(3+i)}{2 \cdot 5 \cdot 10} \\
 &= \frac{5(10i)}{100} = i/2
 \end{aligned}$$

c)

$$\begin{aligned}
 (-1+i)^{50} &= \left(\sqrt{2}e^{i\frac{3\pi}{4}}\right)^{50} \\
 &= 2^{25}e^{i\frac{75\pi}{2}} \\
 &= 2^{25}e^{i\left(\frac{3\pi}{2}+36\pi\right)} \\
 &= 2^{25}(-i) = -2^{25}i
 \end{aligned}$$

Solution 5.5a) Let $z_1 = |z_1|e^{i\theta_1}$, $z_2 = |z_2|e^{i\theta_2}$,

$$\left|\frac{z_1}{z_2}\right| = \left|\frac{|z_1|e^{i\theta_1}}{|z_2|e^{i\theta_2}}\right| = \left|\frac{|z_1|}{|z_2|}e^{i(\theta_1-\theta_2)}\right| = \frac{|z_1|}{|z_2|}$$

Alternatively,

$$\left|\frac{z_1}{z_2}\right|^2 = \left(\frac{z_1}{z_2}\right)\overline{\left(\frac{z_1}{z_2}\right)} = \left(\frac{z_1}{z_2}\right)\left(\frac{\bar{z}_1}{\bar{z}_2}\right) = \frac{z_1\bar{z}_1}{z_2\bar{z}_2} = \frac{|z_1|^2}{|z_2|^2}$$

Since $\left|\frac{z_1}{z_2}\right| > 0$, $|z_1| > 0$, $|z_2| > 0$, we have that

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$$

b) Let $z = re^{i\theta}$. Then

$$z^n = r^n e^{in\theta} = r^n e^{-in\theta} = (re^{-i\theta})^n = (\bar{z})^n$$

c) If $|z| = 1$, then $z = e^{i\theta}$. So

$$\frac{1}{z} = \frac{1}{e^{i\theta}} = e^{-i\theta} = \bar{z}$$

d) Let $z = x + iy$. Then

$$\begin{aligned}(\bar{z})^2 &= (x - iy)^2 = x^2 - y^2 - 2ixy \\ z^2 &= (x + iy)^2 = x^2 - y^2 + 2ixy.\end{aligned}$$

If $(\bar{z})^2 = z^2$, then $x^2 - y^2 - 2ixy = x^2 - y^2 + 2ixy$, therefore the condition is that $-2ixy = 2ixy$, that is, $xy = 0$.

Hence, $(\bar{z})^2 = z^2$ only when z is either real or pure imaginary.

Solution 5.6 If $z^3 = -i$, then

$$z = (-i)^{1/3} = \left(e^{i\frac{3\pi}{2} + 2k\pi}\right)^{1/3}$$

$$k = 0 \Rightarrow \left(e^{i\frac{3\pi}{2}}\right)^{1/3} = e^{i\frac{\pi}{2}} = i$$

$$k = 1 \Rightarrow \left(e^{i\frac{7\pi}{2}}\right)^{1/3} = e^{i\frac{7\pi}{6}} = \cos \frac{7\pi}{6} + i \sin \frac{7\pi}{6} = -\frac{\sqrt{3}}{2} - \frac{1}{2}i$$

$$k = 2 \Rightarrow \left(e^{i\frac{11\pi}{2}}\right)^{1/3} = e^{i\frac{11\pi}{6}} = \cos \frac{11\pi}{6} + i \sin \frac{11\pi}{6} = \frac{\sqrt{3}}{2} - \frac{1}{2}i$$

Chapter 6

Eigen-analysis

6.1 Eigenvalues and eigenvectors

Definition 6.1 *Let A be an $n \times n$ matrix. A number λ and a vector \mathbf{x} are called an eigenvalue eigenvector pair if*

(1) $\mathbf{x} \neq \mathbf{0}$

(2) $A\mathbf{x} = \lambda\mathbf{x}$

In other words, the action of A on the vector \mathbf{x} is to stretch or shrink it by an amount λ without changing its direction. We say λ is an eigenvalue of A if there exists a vector \mathbf{x} so that λ and \mathbf{x} are an eigenvalue eigenvector pair. Notice that we do not allow $\mathbf{x} = \mathbf{0}$. If we did, any number λ would be an eigenvalue. However we do allow $\lambda = 0$. Saying that 0 is an eigenvalue of A means that there is a non-zero solution \mathbf{x} (the eigenvector) of $A\mathbf{x} = 0\mathbf{x} = \mathbf{0}$. So we see that 0 is an eigenvalue of A precisely when A is not invertible.

Let's look at some examples. Consider first the matrix of reflection about a line making an angle of θ with the x axis shown in Figure 6.1. Let \mathbf{x} be any vector that lies along the line. Then the reflection doesn't affect \mathbf{x} . This means that $R\mathbf{x} = \mathbf{x}$. In other words, \mathbf{x} is an eigenvector with eigenvalue 1. On the other hand, suppose that \mathbf{y} is a vector at right angles to the line. Then the reflection flips \mathbf{y} into minus itself. So $R\mathbf{y} = -\mathbf{y}$. In other words, \mathbf{y} is an eigenvector with eigenvalue -1 . If we take any other vector and reflect it, we don't end up with a vector that lies on the same line as the original vector. Thus there are no further eigenvectors or eigenvalues.

An important point to notice is that the eigenvector is not uniquely determined. The vector \mathbf{x} could be *any* vector along the line, and \mathbf{y} could be *any* vector orthogonal to the line. In fact, if we go back to the original definition of eigenvalue and eigenvector we can see that if λ and \mathbf{x} are an eigenvalue eigenvector pair, then so are λ and $s\mathbf{x}$ for any non-zero number s , since $s\mathbf{x} \neq \mathbf{0}$ and $As\mathbf{x} = sA\mathbf{x} = s\lambda\mathbf{x} = \lambda s\mathbf{x}$. So the important thing about an eigenvector is its

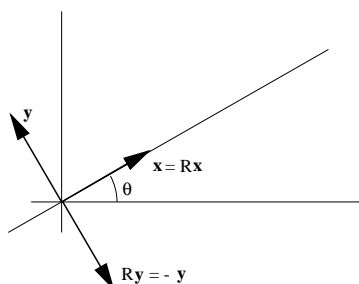


Figure 6.1: Eigenvalues and eigenvectors of a 2D reflection.

direction, not its length. However, there is no such ambiguity in the definition of the eigenvalue. The reflection matrix has exactly two eigenvalues: 1 and -1 .

In some sense, the reflection matrix R illustrates the most satisfactory situation. R is a 2×2 matrix with two distinct eigenvalues. The corresponding eigenvectors \mathbf{x} and \mathbf{y} are linearly independent (in fact they are orthogonal) and form a basis for two dimensional space. It will be important in applications to determine whether or not there exists a basis of eigenvectors of a given matrix. In this example, \mathbf{x} and \mathbf{y} are a basis of eigenvectors of R .

As our next example, consider the identity matrix I . Since the identity matrix doesn't change a vector, we have $I\mathbf{x} = \mathbf{x}$ for any vector \mathbf{x} . Thus any vector \mathbf{x} is an eigenvector of I with eigenvalue 1. This example shows that a given eigenvalue may have many eigenvectors associated with it. However, in this example, there still exists a basis of eigenvectors: any basis at all is a basis of eigenvectors of I .

Next we will consider the rotation matrix ... and run into trouble. Suppose R is the matrix of rotation by $\pi/4$ (*i.e.* 45°). Then $R\mathbf{x}$ is never in the same direction as \mathbf{x} , since R changes the direction of \mathbf{x} by $\pi/4$. So R has no eigenvalues and no eigenvectors. This unfortunate state of affairs will cause us to make a considerable detour into the theory of complex numbers. It turns out that if we work with complex numbers rather than real numbers, then the rotation matrix has eigenvalues too.

6.1.1 Computing the eigenvalues and eigenvectors

We now consider the problem of finding all eigenvalue eigenvector pairs for a given $n \times n$ matrix A . To start, suppose someone tells you that a particular value λ is an eigenvalue of A . How can you find the corresponding eigenvector \mathbf{x} ? This amounts to solving the equation $A\mathbf{x} = \lambda\mathbf{x}$ for \mathbf{x} . This can be rewritten

$$(A - \lambda I)\mathbf{x} = \mathbf{0},$$

where I denotes the identity matrix. In other words \mathbf{x} is a non-zero solution to a homogeneous equation. It can be found by Gaussian elimination.

For example, suppose you know that 4 is an eigenvalue of

$$\begin{bmatrix} 3 & -6 & -7 \\ 1 & 8 & 5 \\ -1 & -2 & 1 \end{bmatrix}.$$

To find the corresponding eigenvector, we must solve

$$\left(\begin{bmatrix} 3 & -6 & -7 \\ 1 & 8 & 5 \\ -1 & -2 & 1 \end{bmatrix} - 4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This can be written

$$\begin{bmatrix} -1 & -6 & -7 \\ 1 & 4 & 5 \\ -1 & -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

To solve this we reduce the matrix. This yields

$$\begin{bmatrix} -1 & -6 & -7 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

The fact that the rank of this matrix is less than 3 confirms that 4 is indeed an eigenvalue. If the rank of the matrix were 3 then the only solution to the equation would be $\mathbf{0}$ which is not a valid eigenvector.

Taking $x_3 = s$ as a parameter, we find that $x_2 = -s$ and $x_1 = -s$. Thus

$$\mathbf{x} = s \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

is an eigenvector (for any non-zero choice of s). In particular, we could take $s = 1$. Then

$$\mathbf{x} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.$$

When doing calculations by hand, it makes sense to take the scalar multiple s in the eigenvector calculation so that it simplifies the form of the eigenvector (clears common denominators in the components, for example). Eigenvectors computed in MATLAB are normalized (scaled so that they have length 1).

Now that we have a method for finding the eigenvectors once we know the eigenvalues, the natural question is: Is there a way to determine the eigenvalues without knowing the eigenvectors? This is where determinants come in. The number λ is an eigenvalue if there is some non-zero solution \mathbf{x} to the equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$. In other words, λ is an eigenvalue if the matrix $(A - \lambda I)$ is not invertible. This happens precisely when $\det(A - \lambda I) = 0$.

This gives us a method for finding the eigenvalues. Compute $\det(A - \lambda I)$. This will be a polynomial in λ . The eigenvalues will be exactly the values of λ that make this polynomial zero, i.e., the roots of the polynomial.

So here is the algorithm for finding the eigenvalues and eigenvectors:

- (1) Compute $\det(A - \lambda I)$ and find the values of λ for which it is zero. These are the eigenvalues.
- (2) For each eigenvalue, find the non-zero solutions to $(A - \lambda I)\mathbf{x} = \mathbf{0}$. These are the eigenvectors.

I should mention that this is actually only a practical way to find eigenvalues when the matrix is small. Finding eigenvalues of large matrices is an important problem and many efficient methods have been developed for use on computers.

Example 6.1 Find the eigenvalues and eigenvectors of

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

First we compute

$$\begin{aligned} \det(A - \lambda I) &= (2 - \lambda)(2 - \lambda) - 1 \\ &= \lambda^2 - 4\lambda + 3 \end{aligned}$$

We can find the roots of this polynomial using the quadratic formula or by factoring it by inspection. We get

$$\lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3),$$

so the eigenvalues are 1 and 3. Now we find the eigenvector for $\lambda = 1$. We must solve $(A - I)\mathbf{x} = \mathbf{0}$. The matrix for this homogeneous system of equations is

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

Reducing this matrix yields

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

so an eigenvector is

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Next we find the eigenvector for $\lambda = 3$. We must solve $(A - 3I)\mathbf{x} = \mathbf{0}$. The matrix for this homogeneous system of equations is

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

Reducing this matrix yields

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

so an eigenvector is

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

Example 6.2 *Let us find the eigenvalues and eigenvectors of*

$$A = \begin{bmatrix} 3 & -6 & -7 \\ 1 & 8 & 5 \\ -1 & -2 & 1 \end{bmatrix}.$$

First we compute

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 3 - \lambda & -6 & -7 \\ 1 & 8 - \lambda & 5 \\ -1 & -2 & 1 - \lambda \end{bmatrix} \\ &= (3 - \lambda)((8 - \lambda)(1 - \lambda) + 10) + 6((1 - \lambda) + 10) - 7(-2 + (8 - \lambda)) \\ &= -\lambda^3 + 12\lambda^2 - 44\lambda + 48 \end{aligned}$$

It is not always easy to find the zeros of a polynomial of degree 3. However, if we already know one solution, we can find the other two. Sometimes, one can find one solution by guessing. In this case we already know that 4 is a solution (since this is the same matrix that appeared in the example in the last section). We can check this:

$$-64 + 12 \times 16 - 44 \times 4 - 48 = 0$$

This means that $\lambda^3 + 12\lambda^2 - 44\lambda + 48$ can be factored as $-\lambda^3 + 12\lambda^2 - 44\lambda + 48 = (\lambda - 4)q(\lambda)$, where $q(\lambda)$ is a second degree polynomial. To find $q(\lambda)$ we can use long division of polynomials.

$$\begin{array}{r} \lambda - 4 \quad) \quad \begin{array}{r} -\lambda^3 + 8\lambda^2 - 12\lambda + 48 \\ -\lambda^3 + 12\lambda^2 - 44\lambda + 48 \\ \hline 8\lambda^2 - 44\lambda \\ 8\lambda^2 - 32\lambda \\ \hline -12\lambda + 48 \\ -12\lambda + 48 \\ \hline 0 \end{array} \end{array}$$

This yields $q(\lambda) = -\lambda^2 + 8\lambda - 12$. This can be factored using the quadratic formula (or by inspection) as $q(\lambda) = -(\lambda - 2)(\lambda - 6)$. So we conclude

$$-\lambda^3 + 12\lambda^2 - 44\lambda + 48 = -(\lambda - 4)(\lambda - 2)(\lambda - 6)$$

and the eigenvalues are 2, 4 and 6. Now we find the eigenvector for $\lambda = 2$. We must solve $(A - 2I)\mathbf{x} = \mathbf{0}$. The matrix for this homogeneous system of equations is

$$\begin{bmatrix} 1 & -6 & -7 \\ 1 & 6 & 5 \\ -1 & -2 & -1 \end{bmatrix}$$

Reducing this matrix yields

$$\begin{bmatrix} 1 & -6 & -7 \\ 0 & -8 & -8 \\ 0 & 0 & 0 \end{bmatrix}$$

so an eigenvector is

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Next we find the eigenvector for $\lambda = 4$. We must solve $(A - 4I)\mathbf{x} = \mathbf{0}$. The matrix for this homogeneous system of equations is

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

Reducing this matrix yields

$$\begin{bmatrix} -1 & -6 & -7 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

so an eigenvector is

$$\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

Finally we find the eigenvector for $\lambda = 6$. We must solve $(A - 6I)\mathbf{x} = \mathbf{0}$. The matrix for this homogeneous system of equations is

$$\begin{bmatrix} -3 & -6 & -7 \\ 1 & 2 & 5 \\ -1 & -2 & -5 \end{bmatrix}$$

Reducing this matrix yields

$$\begin{bmatrix} -3 & -6 & -7 \\ 0 & 0 & 8 \\ 0 & 0 & 0 \end{bmatrix}$$

so an eigenvector is

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

Example 6.3 (Repeated Eigenvalues) Find the eigenvalues and eigenvectors of

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

Consider

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

In this case it makes sense to expand along the last column. This yields

$$\det(A - \lambda I) = 0 - 0 + (1 - \lambda)(1 - \lambda)(2 - \lambda) = (1 - \lambda)^2(2 - \lambda)$$

This is already factored, so the zeros are $\lambda = 1$ and $\lambda = 2$. Notice that the factor $(1 - \lambda)$ occurs to the second power. In this situation there are fewer distinct eigenvalues than we expect. Lets compute the eigenvectors. To find the eigenvectors for $\lambda = 1$ we must solve the homogeneous equation with matrix $A - I$, i.e.,

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

This reduces to

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

and we find that there are two parameters in the solution. The set of solutions in parametric form is

$$s \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We can find two linearly independent solutions by setting $s = 1, t = 0$ and $s = 0, t = 1$. This gives

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

To find the eigenvectors for $\lambda = 2$ we must solve the homogeneous equation with matrix $A - 2I$, i.e.,

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

This reduces to

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

and we find that the set of solutions in parametric form is

$$s \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Setting $s = 1$ gives the eigenvector

$$\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

In this 3×3 example, even though there are only two distinct eigenvalues, 1 and 2, there are still three independent eigenvectors (i.e., a basis), because the eigenvalue 1 has two independent eigenvectors associated to it.

Example 6.4 (Repeated Eigenvalues with “Missing” Eigenvectors) *Find the eigenvalues and eigenvectors of*

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

Here

$$\det(A - \lambda I) = (\lambda - 2)^2$$

so there is only one eigenvalue $\lambda = 2$. To find the eigenvectors, we must solve the homogeneous system with matrix

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

The solutions are

$$s \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

so there is only one eigenvector direction. So here is a matrix that does not have a basis of eigenvectors. Matrices like this, that have too few eigenvectors, will not be studied further in this course, but they do occur in applications.

6.1.2 Complex eigenvalues and eigenvectors

Since eigenvalues are found as roots of polynomials, we now see that they can be complex. A discussion of complex eigenvalues and eigenvectors is given below.

Example 6.5 *Lets consider the matrix of rotation by $\pi/2$. This is the matrix*

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

We compute

$$\det(A - \lambda I) = \det \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} = \lambda^2 + 1$$

The roots are $\pm i$ so the eigenvalues are i and $-i$. Now we compute the eigenvector corresponding to the eigenvalue i . We must solve the homogeneous equation with matrix

$$\begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix}$$

Notice that we will have to do complex arithmetic to achieve this, since the matrix now has complex entries. To reduce this matrix we have to add i times the first row to the second row. This gives

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

So if we let the $x_2 = s$, then $-ix_1 + s = 0$, or $x_1 = -is$. So the solution is

$$s \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

and we may choose $s = 1$. Lets check that this is really an eigenvector:

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

To find the other eigenvector we can use a trick. Suppose that the original matrix A has only real entries. This will always be the case in our examples. Suppose that A has a complex eigenvalue eigenvector pair λ and \mathbf{x} . Then $A\mathbf{x} = \lambda\mathbf{x}$. Taking the complex conjugate of this equation, we obtain $\bar{A}\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}$. (Here conjugating a matrix or a vector just means conjugating each entry). Now, since A has real entries, $\bar{A} = A$. Hence $A\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}$. In other words $\bar{\lambda}$ is an eigenvalue with eigenvector $\bar{\mathbf{x}}$. In the present example, we already know that $\bar{i} = -i$ is an eigenvalue. But now we don't have to compute the eigenvector that goes along with it. It is simply the conjugate of the one we already computed. So the eigenvector corresponding to $-i$ is

$$\begin{bmatrix} i \\ 1 \end{bmatrix}$$

The eigenvalues of A are the zeros or roots of the polynomial $\det(A - \lambda I)$. If we use complex numbers then $\det(A - \lambda I)$ can be completely factored, i.e.,

$$\det(A - \lambda I) = \pm(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

Finding the roots may be difficult. However for 2×2 matrices we may use the quadratic formula.

If all the roots are distinct (i.e., $\lambda_i \neq \lambda_j$ for $i \neq j$) then the corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ are linearly independent (I didn't show you why this is true, so I'm just asking you to believe it!) and therefore form a basis.

If there are repeated roots, then there are fewer than n distinct eigenvalues. In this situation, it might happen that there are not enough eigenvectors to

form a basis. However it also might happen that more than one eigenvector associated to a given eigenvalue, so that in the end there are enough eigenvectors to form a basis. Compare Examples 6.3 and 6.4 from earlier, where we saw that either situation can occur. Unfortunately, the only way we have to find out is to try to compute them all.

6.1.3 MATLAB

When applied to a square matrix \mathbf{A} , `eig(A)` will return the eigenvalues and the eigenvectors of \mathbf{A} . To use this command enter the following in MATLAB:

```
>> [P,D] = eig(A)
```

What will be returned is a matrix \mathbf{P} with the normalized (unit length) eigenvectors of \mathbf{A} in its columns, and a matrix \mathbf{D} with the eigenvalues of \mathbf{A} along its diagonal. The eigenvalue corresponding to the i^{th} column of \mathbf{P} is found in the (i, i) position of \mathbf{D} . Using the `eig` command above will return complex eigenvalues and eigenvectors when present.

Example 6.6 Consider A

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 5 \\ 6 & 3 & 9 \\ 2 & 7 & 8 \end{bmatrix} \quad (6.1)$$

We can enter \mathbf{A} into MATLAB and find its eigenvectors and eigenvalues with the following commands:

```
>> A=[1 4 5; 6 3 9; 2 7 8];
>> [P,D] = eig(A)
```

$\mathbf{P} =$

```
-0.3919   -0.5895    0.2238
-0.6401   -0.5446   -0.8511
-0.6609    0.5966    0.4750
```

$\mathbf{D} =$

```
15.9657         0         0
         0   -0.3653         0
         0         0   -3.6004
```

These results tell us that A has eigenvectors $\{v_1, v_2, v_3\}$ and corresponding eigenvalues $\{\lambda_1, \lambda_2, \lambda_3\}$ as follows:

$$\{v_1, v_2, v_3\} \approx \left\{ \begin{pmatrix} -0.3919 \\ -0.6401 \\ -0.6609 \end{pmatrix}, \begin{pmatrix} -0.5895 \\ -0.5446 \\ 0.5966 \end{pmatrix}, \begin{pmatrix} 0.2238 \\ -0.8511 \\ 0.4750 \end{pmatrix} \right\}$$

$$\{\lambda_1, \lambda_2, \lambda_3\} \approx \{15.9657, -0.3653, -3.6004\}$$

6.1.4 Problems

Problem 6.1 Show that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are eigenvectors for the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. What are the corresponding eigenvalues?

Problem 6.2 Suppose P is a projection matrix. What are the eigenvalues and eigenvectors of P ?

Problem 6.3 Find the eigenvalues and eigenvectors for

$$a) \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix} \quad b) \begin{bmatrix} -2 & -8 \\ 4 & 10 \end{bmatrix} \quad c) \begin{bmatrix} 29 & -10 \\ 105 & -36 \end{bmatrix} \quad d) \begin{bmatrix} -9 & -14 \\ 7 & 12 \end{bmatrix}$$

Problem 6.4 Find the eigenvalues and the corresponding eigenvectors of the matrix

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

Problem 6.5 Find the eigenvalues and eigenvectors for

$$a) \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 2 \\ 2 & 0 & 2 \end{bmatrix} \quad b) \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{bmatrix} \quad c) \begin{bmatrix} 7 & -9 & -15 \\ 0 & 4 & 0 \\ 3 & -9 & -11 \end{bmatrix} \quad d) \begin{bmatrix} 31 & -100 & 70 \\ 18 & -59 & 42 \\ 12 & -40 & 29 \end{bmatrix}$$

Problem 6.6 Let P be a 2×2 transitional probability matrix in the form

$$\begin{bmatrix} p_{11} & 1 - p_{22} \\ 1 - p_{11} & p_{22} \end{bmatrix}.$$

Prove that one of the eigenvalues of P must be 1, and another one must be in the interval $[-1, 1]$. (Hint: Let $c = p_{11} + p_{22}$.)

Problem 6.7 Find the eigenvalues and the corresponding eigenvectors of the following matrix.

$$A = \begin{bmatrix} 0 & 1 & -1 \\ 5 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

Problem 6.8 Find the eigenvalues and the corresponding eigenvectors of the following matrix.

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

Problem 6.9 Is there a rank two matrix M , such that vectors $\vec{v}_{\mu_1} = [1, 2, 3]^T$ and $\vec{v}_{\mu_2} = [3, 2, 1]^T$ are eigenvectors of M , both corresponding to the same eigenvalue $\mu_1 = \mu_2 = -1$? If your answer is yes then find such a matrix, and if your answer is no, then justify your answer.

Problem 6.10 Find the eigenvalues and the corresponding eigenvectors of the matrix

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

Problem 6.11 Given a 2×2 matrix

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

- a, Find values for a and b that $A^2 = A$.
 b, Find values for a and b that $A^3 = A$ and $A^4 \neq A$.
 c, Find values for a and b that $A^8 \neq A$ and $A^9 = A$.

Problem 6.12 Like in the previous question, given a 2×2 matrix

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

- a, Find values for a and b that the two eigenvalues of A are $\mu_1 = 2 + i$ and $\mu_2 = 2 - i$.
 b, Find the eigenvectors of A in the previous question. (where the two eigenvalues are $2 \pm i$)

6.2 Eigenanalysis simplifies matrix powers

In the previous section we learnt how to find eigenvalues and eigenvectors of a matrix, including the case when they are complex. There are two main uses of this eigenanalysis: efficiently computing powers of a matrix (studied in this section) and in the solution of differential equations considered in Section 6.3 below.

Recall that in the random walk application in Section 4.3 we were interested in high powers of a matrix, specifically

$$\lim_{n \rightarrow \infty} P^n \mathbf{x}^{(0)}$$

where P is the matrix of transition probabilities and $\mathbf{x}^{(0)}$ is the column vector of initial probabilities. We will explore the use of eigenanalysis to simplify our understanding of these kind of problems in two examples below.

Example 6.7 We consider again the sorcerers' duel in Example 4.1. We will consider what happens if the duel is allowed to continue without limit until there is a winner. Rather than compute

$$\lim_{n \rightarrow \infty} P^n \mathbf{x}^{(0)}$$

numerically in MATLAB as was done in Section 4.3 we will use an eigenanalysis of P to understand this limit. The transition matrix for this problem is

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

with initial state $\mathbf{x}^{(0)} = (1, 0, 0, 0)^T$. The eigenanalysis of P is summarized below:

$$\begin{aligned} \lambda_1 &= 1 & \mathbf{k}_1 &= (0, 0, 1, 0)^T \\ \lambda_2 &= 1 & \mathbf{k}_2 &= (0, 0, 0, 1)^T \\ \lambda_3 &= \frac{1}{\sqrt{3}} & \mathbf{k}_3 &= (1 - \sqrt{3}, \frac{2}{\sqrt{3}}(1 - \sqrt{3}), \frac{1}{\sqrt{3}}, 1)^T \\ \lambda_4 &= -\frac{1}{\sqrt{3}} & \mathbf{k}_4 &= (1 + \sqrt{3}, -\frac{2}{\sqrt{3}}(1 + \sqrt{3}), -\frac{1}{\sqrt{3}}, 1)^T \end{aligned}$$

Note that $\lambda = 1$ is a repeated eigenvalue but there there is still a basis of eigenvectors (the set of eigenvectors associated with $\lambda = 1$ is two-dimensional). We can write

$$\mathbf{x}^{(0)} = c_1 \mathbf{k}_1 + c_2 \mathbf{k}_2 + c_3 \mathbf{k}_3 + c_4 \mathbf{k}_4 \quad (6.2)$$

for some coefficients c_1, c_2, c_3 and c_4 uniquely determined. Equation (6.2) can be written in matrix-vector form

$$T\mathbf{c} = \mathbf{x}^{(0)} \quad (6.3)$$

where $\mathbf{c} = (c_1, c_2, c_3, c_4)^T$ and T is the 4×4 matrix with eigenvectors $\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3$ and \mathbf{k}_4 in its columns. Solving (6.3) (I used MATLAB) gives $c_1 = 1/2, c_2 = 1/2, c_3 \approx -0.6830$ and $c_4 \approx 0.1830$. With these values of \mathbf{c} (6.2) is a representation of $\mathbf{x}^{(0)}$ as a linear combination of eigenvectors of P . This makes working out later states $\mathbf{x}^{(n)}$ easy, as shown below.

$$\begin{aligned} \mathbf{x}^{(1)} &= P\mathbf{x}^{(0)} = P(c_1 \mathbf{k}_1 + c_2 \mathbf{k}_2 + c_3 \mathbf{k}_3 + c_4 \mathbf{k}_4) \\ &= c_1 \lambda_1 \mathbf{k}_1 + c_2 \lambda_2 \mathbf{k}_2 + c_3 \lambda_3 \mathbf{k}_3 + c_4 \lambda_4 \mathbf{k}_4 \\ &= c_1 \mathbf{k}_1 + c_2 \mathbf{k}_2 + \frac{1}{\sqrt{3}} c_3 \mathbf{k}_3 - \frac{1}{\sqrt{3}} c_4 \mathbf{k}_4 \end{aligned}$$

where in the middle line we have remembered that the \mathbf{k}_i vectors are eigenvectors, so $P\mathbf{k}_i = \lambda \mathbf{k}_i$ for $i = 1, 2, 3, 4$. Similarly

$$\mathbf{x}^{(2)} = P\mathbf{x}^{(1)} = P^2 \mathbf{x}^{(0)} = c_1 \mathbf{k}_1 + c_2 \mathbf{k}_2 + \frac{1}{3} c_3 \mathbf{k}_3 + \frac{1}{3} c_4 \mathbf{k}_4$$

and

$$\mathbf{x}^{(n)} = P^n \mathbf{x}^{(0)} = c_1 \mathbf{k}_1 + c_2 \mathbf{k}_2 + \left(\frac{1}{\sqrt{3}}\right)^n c_3 \mathbf{k}_3 + \left(-\frac{1}{\sqrt{3}}\right)^n c_4 \mathbf{k}_4 \quad (6.4)$$

This formula (6.4) is a simple formula for the state at any time n that does not involve much computational work. In addition, it is easy to see from this formula that

$$\lim_{n \rightarrow \infty} \mathbf{x}^{(n)} = (0, 0, 1/2, 1/2)^T$$

as found numerically in Example 4.1

Example 6.8 *The weather in Vancouver is either good, average or bad on any given day. If the weather is good on any day, there is a 60% chance the weather will be good, 30% chance average, and 10% bad on the next day. If the weather is average, then on the next day there is a 40% chance of good, 30% of average and 30% of bad weather. If the weather is bad then on the next day there is a 40% chance of good, 50% of average and 10% of bad. If the weather is good today, what will the weather be like a long time from now? We number the states*

- 1) good
- 2) average
- 3) bad

The corresponding transition matrix is

$$P = \frac{1}{10} \begin{bmatrix} 6 & 4 & 4 \\ 3 & 3 & 5 \\ 1 & 3 & 1 \end{bmatrix}$$

The initial state is $\mathbf{x}^{(0)} = [1, 0, 0]^T$. The eigenanalysis of P is summarized below:

$$\begin{aligned} \lambda_1 &= 1, & \mathbf{k}_1 &= [1/2, 1/3, 1/6]^T \\ \lambda_2 &= 0.2, & \mathbf{k}_2 &= [-2, 1, 1]^T \\ \lambda_3 &= -0.2, & \mathbf{k}_3 &= [0, -1, 1]^T \end{aligned}$$

As above, we put the eigenvectors into the columns of a matrix T and solve

$$T\mathbf{c} = (1, 0, 0)^T$$

for $\mathbf{c} = (1, -1/4, 1/12)$. The right hand side $(1, 0, 0)^T$ of the system above corresponds to the initial state $\mathbf{x}^{(0)}$ of good weather. As before, this gives us the representation of the initial state as a linear combination of the eigenvectors:

$$\mathbf{x}^{(0)} = \mathbf{k}_1 - \frac{1}{4}\mathbf{k}_2 + \frac{1}{12}\mathbf{k}_3.$$

Again, we see that multiplying by P in this representation leads to an easy formula since \mathbf{k}_i are eigenvectors:

$$\mathbf{x}^{(n)} = P^n \mathbf{x}^{(0)} = \mathbf{k}_1 - \frac{1}{4}(0.2)^n \mathbf{k}_2 + \frac{1}{12}(-0.2)^n \mathbf{k}_3$$

Note that after a long time ($n \rightarrow \infty$) the second and third terms above tend to zero, so

$$\lim_{n \rightarrow \infty} \mathbf{x}^{(n)} = \mathbf{k}_1 = (1/2, 1/3, 1/6)$$

so after a long time after the first nice day, the weather will have a 1/2 chance of being good, 1/3 average and 1/6 bad.

Let us consider the example above a bit more closely. Intuitively, the weather after a long time should not depend on what it was like the day you started. In fact, we can show that

$$\lim_{n \rightarrow \infty} \mathbf{x}^{(n)} = \mathbf{k}_1 \quad (6.5)$$

for any starting probability $x^{(0)}$ (for which the entries must be non-negative and sum to one).

To show this, we have to show that writing

$$\mathbf{x}^{(0)} = c_1 \mathbf{k}_1 + c_2 \mathbf{k}_2 + c_3 \mathbf{k}_3 \quad (6.6)$$

always gives $c_1 = 1$ no matter what $\mathbf{x}^{(0)}$ is. Note that the entries of \mathbf{k}_2 and \mathbf{k}_3 sum to zero and the entries of \mathbf{k}_1 and $x^{(0)}$ sum to one. So by summing the entries of (6.6) we see that $c_1 = 1$ which shows that (6.5) is true for any starting probability $x^{(0)}$ as our intuition predicted. The probability \mathbf{k}_1 that all initial states tend to is called an *equilibrium probability*. In some cases described in the theorem below you can guarantee the existence of an equilibrium probability.

Theorem 6.1 *If P is a transition matrix (non-negative entries with all columns summing to one) that in addition has all positive entries then P has an eigenvalue 1 with a single eigenvector \mathbf{k}_1 that can be chosen to be a probability vector. All other eigenvalues λ satisfy $|\lambda| < 1$ with eigenvectors with components that sum to zero. Thus,*

$$\lim_{n \rightarrow \infty} \mathbf{x}^{(n)} = \mathbf{k}_1$$

for any $x^{(0)}$. That is, \mathbf{k}_1 is an equilibrium probability.

Note that the Example 6.8 the transition matrix satisfied the conditions of the theorem and had an equilibrium probability. The transition matrix of Example 6.7 did not satisfy the conditions of the theorem and does not have an equilibrium probability (depending on the initial state, you can tend to different fractions of times that each sorcerer wins the duel).

We can summarize the process used to analyze these examples: writing a vector \mathbf{x} (the initial probability in our examples) as a linear combination of eigenvectors of a matrix A (the transition matrix for our examples) then easily writing $A^n \mathbf{x}$ as a linear combination of eigenvectors. This process can be summarized in matrix-vector notation as Diagonalization discussed in more detail in Section 6.5.1.

6.2.1 Problems

Problem 6.13 Find the eigenvalues and corresponding eigenvectors of the stochastic matrix P below. Use the eigenvectors and eigenvalues to describe

$$\lim_{n \rightarrow \infty} P^n x^{(0)},$$

where $x^{(0)} = [1, 0, 0]^T$, and

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

Problem 6.14 What is the necessary condition on a and b , for P having an equilibrium probability? Find the equilibrium probability vector (if it exists) for $a = b = 1/4$.

$$P = \begin{bmatrix} \frac{1}{4} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} & b \\ \frac{1}{4} & \frac{1}{3} & a \end{bmatrix}$$

6.3 Systems of linear differential equations

Consider the system of differential equations

$$\begin{aligned} y_1'(t) &= a_{1,1}y_1(t) + a_{1,2}y_2(t) \\ y_2'(t) &= a_{2,1}y_1(t) + a_{2,2}y_2(t) \end{aligned} \quad (6.7)$$

This system of equations describes a situation where we have two quantities y_1 and y_2 , where the rate of change of each one of the quantities depends on the values of both. We can rewrite this as a matrix equation. Let $\mathbf{y}(t)$ be the vector

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix},$$

and define the derivative of a vector to be the vector of derivatives, i.e.,

$$\mathbf{y}'(t) = \begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix}.$$

Define A to be the matrix

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

Then the system of equations (6.7) can be rewritten

$$\mathbf{y}'(t) = A\mathbf{y}.$$

A general system of linear equations has this form, except $\mathbf{y}(t)$ is an n -dimensional vector and A is an $n \times n$ matrix. How can we find solutions to such

a system of equations? Taking a hint from the scalar case, we can try to find solutions of the form

$$\mathbf{y}(t) = e^{\lambda t} \mathbf{x}$$

where \mathbf{x} is a fixed vector (not depending on t). With this definition

$$\mathbf{y}'(t) = \lambda e^{\lambda t} \mathbf{x}$$

so that $\mathbf{y}' = A\mathbf{y}$ whenever

$$\lambda e^{\lambda t} \mathbf{x} = A e^{\lambda t} \mathbf{x} = e^{\lambda t} A \mathbf{x}$$

Dividing by $e^{\lambda t}$, this condition becomes

$$\lambda \mathbf{x} = A \mathbf{x}.$$

In other words, $\mathbf{y}(t) = e^{\lambda t} \mathbf{x}$ is a solution exactly whenever λ and \mathbf{x} are an eigenvalue eigenvector pair for A . So we can find as many solutions as we have eigenvalue eigenvector pairs.

To proceed we first notice that if $\mathbf{y}_1(t)$ and $\mathbf{y}_2(t)$ are two solutions to the equation $\mathbf{y}' = A\mathbf{y}$, then a linear combination $c_1\mathbf{y}_1(t) + c_2\mathbf{y}_2(t)$ is also a solution, since

$$\begin{aligned} \frac{d}{dt} (c_1\mathbf{y}_1(t) + c_2\mathbf{y}_2(t)) &= c_1\mathbf{y}'_1(t) + c_2\mathbf{y}'_2(t) \\ &= c_1A\mathbf{y}_1(t) + c_2A\mathbf{y}_2(t) \\ &= A(c_1\mathbf{y}_1(t) + c_2\mathbf{y}_2(t)) \end{aligned}$$

Notice that we are assuming that the constants c_1 and c_2 do not depend on t . Similarly, if $\mathbf{y}_1(t), \mathbf{y}_2(t), \dots, \mathbf{y}_n(t)$ are n solutions then $c_1\mathbf{y}_1(t) + c_2\mathbf{y}_2(t) + \dots + c_n\mathbf{y}_n(t)$ is a solution for any choice of constants c_1, c_2, \dots, c_n .

Now suppose that A is an $n \times n$ matrix. Suppose that $\lambda_1, \lambda_2, \dots, \lambda_k$ are its eigenvalues with eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$. Then we have that for any choice of constants c_1, c_2, \dots, c_k ,

$$\mathbf{y}(t) = c_1 e^{\lambda_1 t} \mathbf{x}_1 + c_2 e^{\lambda_2 t} \mathbf{x}_2 + \dots + c_k e^{\lambda_k t} \mathbf{x}_k \quad (6.8)$$

is a solution. Have we found all solutions? In other words, could there be a solution of the equation that is not this form, or is every solution of the form (6.8) for some choice of c_1, c_2, \dots, c_k ?

There is a theorem in differential equations that says that given an initial condition \mathbf{x}_0 there is one and only one solution of $\mathbf{y}' = A\mathbf{y}$ satisfying $\mathbf{y}(0) = \mathbf{y}_0$. So our theoretical question above is equivalent to the following quite practical question. Given an initial vector \mathbf{y}_0 , does there exist a solution $\mathbf{y}(t)$ of the form (6.8) whose value at zero is the given initial condition, i.e., $\mathbf{y}(0) = \mathbf{y}_0$?

This will be true if, given any vector \mathbf{x}_0 , one can find c_1, c_2, \dots, c_k so that

$$\mathbf{y}(0) = c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \dots + c_k \mathbf{x}_k = \mathbf{x}_0$$

This is exactly the condition that the eigenvectors form a basis. It turns out that in the “bad” cases where there are not enough eigenvectors of A to form a basis, there are solutions that don’t have the form (6.8).

Now suppose that there are n eigenvectors that do form a basis. How can we actually find the numbers c_1, c_2, \dots, c_n such that

$$c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \dots + c_n \mathbf{x}_n = \mathbf{x}_0?$$

Just notice that this is a system linear equations

$$\left[\begin{array}{c|c|c|c} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{array} \right] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{x}_0$$

so you know what to do.

Example 6.9 Find the general solution to the system of equations

$$\begin{aligned} y_1'(t) &= 2y_1(t) + y_2(t) \\ y_2'(t) &= y_1(t) + 2y_2(t) \end{aligned}$$

This is equivalent to the matrix equation

$$\mathbf{y}'(t) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \mathbf{y}(t)$$

The matrix $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ has eigenvector and eigenvalues $\lambda_1 = 1$, $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\lambda_2 = 3$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The eigenvectors \mathbf{x}_1 and \mathbf{x}_2 form a basis, so the general solution is

$$\mathbf{y}(t) = c_1 e^{\lambda_1 t} \mathbf{x}_1 + c_2 e^{\lambda_2 t} \mathbf{x}_2 = c_1 e^t \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 e^{3t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Example 6.10 Continue Example 6.9 above and find the solution satisfying the initial condition

$$\mathbf{y}(0) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

We have to find constants c_1 and c_2 so that

$$c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

This is the same as solving

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

The solution is

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 3/2 \end{bmatrix}$$

Example 6.11 Now let's do an example where the eigenvalues are complex. Consider the equation

$$\mathbf{y}'(t) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{y}(t).$$

Find the general solution of this differential equation. The matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ has eigenvector and eigenvalues $\lambda_1 = i$, $\mathbf{x}_1 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$ and complex conjugate $\lambda_2 = -i$, $\mathbf{x}_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}$. The eigenvectors \mathbf{x}_1 and \mathbf{x}_2 form a basis, so the general solution is

$$\mathbf{y}(t) = c_1 e^{\lambda_1 t} \mathbf{x}_1 + c_2 e^{\lambda_2 t} \mathbf{x}_2 = c_1 e^{it} \begin{bmatrix} -i \\ 1 \end{bmatrix} + c_2 e^{-it} \begin{bmatrix} i \\ 1 \end{bmatrix}$$

In most applications, the solutions that we are interested in are real. The solution above looks decidedly complex! Remember, however, that the constants c_1 and c_2 can be complex too. Perhaps for special choices of c_1 and c_2 the solution will turn out to be real. This is in fact always true when the original matrix is real. In this case the complex eigenvalues and eigenvectors occur in conjugate pairs. So if

$$\mathbf{y}_1(t) = e^{\lambda t} \mathbf{x}$$

is a solution, then so is

$$\bar{\mathbf{y}}_1(t) = e^{\bar{\lambda} t} \bar{\mathbf{x}}$$

So if we choose $c_1 = a/2$ and $c_2 = a/2$ for a real number a , then

$$\begin{aligned} c_1 e^{\lambda t} \mathbf{x} + c_2 e^{\bar{\lambda} t} \bar{\mathbf{x}} &= a/2 (e^{\lambda t} \mathbf{x} + e^{\bar{\lambda} t} \bar{\mathbf{x}}) \\ &= a \Re(e^{\lambda t} \mathbf{x}) \end{aligned}$$

(here \Re stands for the real part. We used that for a complex number z , $z + \bar{z} = 2\Re z$). Similarly, if we choose $c_1 = a/2i$ and $c_2 = -a/2i$, then

$$\begin{aligned} c_1 e^{\lambda t} \mathbf{x} + c_2 e^{\bar{\lambda} t} \bar{\mathbf{x}} &= a/2i (e^{\lambda t} \mathbf{x} - e^{\bar{\lambda} t} \bar{\mathbf{x}}) \\ &= a \Im(e^{\lambda t} \mathbf{x}) \end{aligned}$$

where $\Im()$ denotes the imaginary part of the argument. The upshot is that the real and imaginary parts of a solution are also solutions. It's sometimes easier to just start with one of the complex solutions and find its real and imaginary parts. This gives us two real solutions to work with. Notice that it doesn't matter which one of the complex solutions we pick. Because they are conjugate, their real parts are the same, and their imaginary parts differ by only a minus sign.

Example 6.12 Continuing Example 6.11 In the example we have

$$\mathbf{y}_1(t) = e^{it} \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} -ie^{it} \\ e^{it} \end{bmatrix} \\
&= \begin{bmatrix} -i(\cos(t) + i\sin(t)) \\ \cos(t) + i\sin(t) \end{bmatrix} \\
&= \begin{bmatrix} -i\cos(t) + \sin(t) \\ \cos(t) + i\sin(t) \end{bmatrix} \\
&= \begin{bmatrix} \sin(t) \\ \cos(t) \end{bmatrix} + i \begin{bmatrix} -\cos(t) \\ \sin(t) \end{bmatrix}
\end{aligned}$$

The real part and imaginary part are

$$\begin{bmatrix} \sin(t) \\ \cos(t) \end{bmatrix}$$

and

$$\begin{bmatrix} -\cos(t) \\ \sin(t) \end{bmatrix}$$

One can check directly that these are solutions to the original equation. The general solution can also be written

$$a_1 \begin{bmatrix} \sin(t) \\ \cos(t) \end{bmatrix} + a_2 \begin{bmatrix} -\cos(t) \\ \sin(t) \end{bmatrix}$$

The advantage of this way of writing the solution is that if we choose a_1 and a_2 to be real the solution is real too.

Example 6.13 Continuing Example 6.12 Now suppose we want to satisfy an initial condition. Let's find the solution $\mathbf{y}(t)$ of the equation that satisfies

$$\mathbf{y}(0) = \begin{bmatrix} 2 \\ -2 \end{bmatrix}.$$

There are two ways to proceed. we could use the complex form of the general solution. Then we must find c_1 and c_2 such that

$$c_1 \begin{bmatrix} -i \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

This amounts to solving

$$\begin{bmatrix} -i & i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

The solution is

$$\begin{aligned}
\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} -i & i \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ -2 \end{bmatrix} \\
&= \frac{1}{-2i} \begin{bmatrix} 1 & -i \\ -1 & -i \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} \\
&= \begin{bmatrix} i + 1/2 \\ i - 1/2 \end{bmatrix}
\end{aligned}$$

So $c_1 = i + 1/2$ and $c_2 = i - 1/2$. If we plug these into the expression for the general solution we get the right answer. However, there is still a fair amount of complex arithmetic needed to show explicitly that the solution is real. It's easier to start with the real solutions. In this approach we must find a_1 and a_2 so that

$$a_1 \begin{bmatrix} \sin(0) \\ \cos(0) \end{bmatrix} + a_2 \begin{bmatrix} -\cos(0) \\ \sin(0) \end{bmatrix} = a_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

Thus $a_1 = a_2 = -2$ so the solution is

$$-2 \begin{bmatrix} \sin(t) \\ \cos(t) \end{bmatrix} + -2 \begin{bmatrix} -\cos(t) \\ \sin(t) \end{bmatrix} = \begin{bmatrix} -2\sin(t) + 2\cos(t) \\ -2\cos(t) - 2\sin(t) \end{bmatrix}$$

Example 6.14 Now let's do an example where the eigenvalues are complex, and have both a real and imaginary part. Let's solve

$$\mathbf{y}'(t) = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \mathbf{y}(t)$$

with initial condition

$$\mathbf{y}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The first step is to find the eigenvalues and eigenvectors. I'll omit the computations. The result is $\lambda_1 = -1 + i$ with eigenvector $\mathbf{x}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ and the complex conjugates $\lambda_2 = -1 - i$ with eigenvector $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$. Thus a solution is

$$\mathbf{y}_1(t) = e^{(-1+i)t} \begin{bmatrix} 1 \\ i \end{bmatrix}$$

To find real solutions we calculate the real and imaginary parts of this.

$$\begin{aligned} \mathbf{y}_1(t) &= \begin{bmatrix} e^{(-1+i)t} \\ ie^{(-1+i)t} \end{bmatrix} \\ &= \begin{bmatrix} e^{-t}e^{it} \\ ie^{-t}e^{it} \end{bmatrix} \\ &= \begin{bmatrix} e^{-t}(\cos(t) + i\sin(t)) \\ ie^{-t}(\cos(t) + i\sin(t)) \end{bmatrix} \\ &= \begin{bmatrix} e^{-t}\cos(t) \\ -e^{-t}\sin(t) \end{bmatrix} + i \begin{bmatrix} e^{-t}\sin(t) \\ e^{-t}\cos(t) \end{bmatrix} \end{aligned}$$

So the general solution can be written

$$a_1 \begin{bmatrix} e^{-t}\cos(t) \\ -e^{-t}\sin(t) \end{bmatrix} + a_2 \begin{bmatrix} e^{-t}\sin(t) \\ e^{-t}\cos(t) \end{bmatrix}$$

To satisfy the initial condition, we need

$$a_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

so that $a_1 = 1$ and $a_2 = 1$. Thus the solution is

$$\mathbf{y}(t) = e^{-t} \begin{bmatrix} \cos(t) + \sin(t) \\ -\sin(t) + \cos(t) \end{bmatrix}$$

6.3.1 Problems

Problem 6.15 Find the general solution to $\mathbf{y}' = \mathbf{A}\mathbf{y}$ when $A = \begin{bmatrix} -2 & -8 \\ 4 & 10 \end{bmatrix}$. (Hint: This matrix appeared in the problems of last chapter). Find the solution satisfying the initial condition $\mathbf{y}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Problem 6.16 Find the general solution to $\mathbf{y}' = \mathbf{A}\mathbf{y}$ when $A = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$. Find both the complex form and the real form. Find the solution satisfying the initial condition $\mathbf{y}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Problem 6.17 Find the general solution to $\mathbf{y}' = \mathbf{A}\mathbf{y}$ when $A = \begin{bmatrix} 6 & 0 & 13 \\ 5 & 1 & 13 \\ -2 & 0 & -4 \end{bmatrix}$. Find both the complex form and the real form. Find the solution satisfying the initial condition $\mathbf{y}(0) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Problem 6.18 Find the general solution of the following system of differential equations:

$$\mathbf{y}'(t) = \begin{bmatrix} 1 & 1 \\ 5 & 1 \end{bmatrix} \mathbf{y}(t)$$

Problem 6.19 Find the solution of the following system of differential equations, that satisfy the initial condition $\mathbf{y}(0) = [0, 1]^T$.

$$\mathbf{y}'(t) = \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix} \mathbf{y}(t)$$

Problem 6.20 Find the real form of the solution of the following system of differential equations, that satisfy the initial condition $\mathbf{y}(0) = [2, -1]^T$.

$$\mathbf{y}'(t) = \begin{bmatrix} -3 & -1 \\ 7 & 1 \end{bmatrix} \mathbf{y}(t)$$

Problem 6.21 Is it true that every 3×3 matrix with real entries always has at least one real eigenvalue? Why?

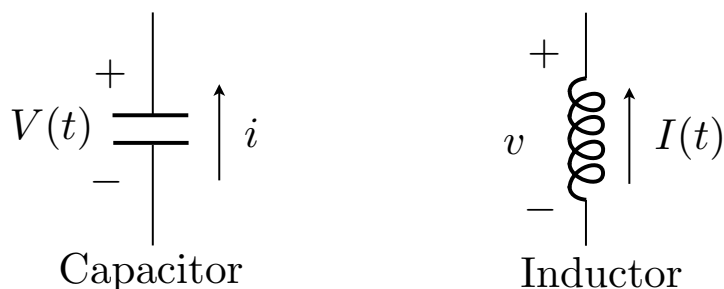


Figure 6.2: Diagrams of capacitors and inductors.

6.4 LCR circuits

6.4.1 Capacitors and inductors

We introduce new elements into our study of electrical circuits: capacitors and inductors. At an instant in time t , a capacitor acts as a voltage source with voltage $V(t)$. At a given instant in time an inductor acts as a current source with current $I(t)$. These circuit elements are shown in Figure 6.2.

The voltage across the capacitor changes proportional to the current i (with direction as in Figure 6.2) through it

$$\frac{dV}{dt} = -\frac{i}{C} \quad (6.9)$$

where the constant of proportionality C is the capacitance of the capacitor with MKS units of Farads. Note that a capacitor with large capacitance can provide more current for the same drop in voltage.

The current through an inductor changes proportional to the voltage v across it (with direction as in Figure 6.2)

$$\frac{dI}{dt} = -\frac{v}{L} \quad (6.10)$$

where the constant of proportionality L is the inductance with MKS units of Henrys.

We consider here at a high level how the behaviour in time of circuits with many capacitors, inductors and resistors. At a given instant in time, capacitors can be treated as a voltage sources and inductors as current sources. To determine how these sources change in time, the current through the capacitors (6.9) and the voltage across the inductors (6.10) is needed. However, determining these from the sources is the *fundamental circuit problem* considered in Section 3.5.3 and Section 4.7.1. Thus, a system of differential equations can be derived for circuit networks with capacitors and inductors by solving the fundamental problem for the circuit and scaling the result by the capacitance and inductance of the elements. Some specific examples are given below.

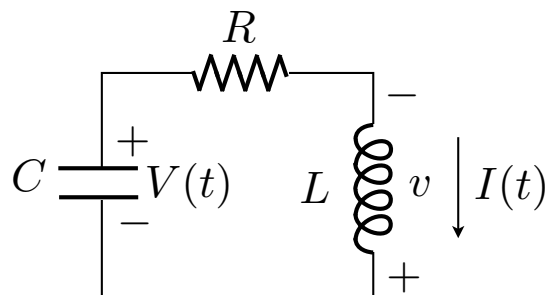


Figure 6.3: A simple LRC circuit.

6.4.2 Differential equations for LCR circuits

Example 6.15 *The simplest example of an LRC circuit is the simple series circuit shown in Figure 6.3. We will derive a system of differential equations for $V(t)$ and $I(t)$ in the circuit and determine what combination of values for the components L , R , and C lead to oscillations. The equation is simple enough that the solution to the fundamental problem can just be written down. The current through the capacitor is I and the voltage v across the inductor is $IR - V$. Using the relationships (6.9, 6.10) we write*

$$\begin{aligned}\frac{dV}{dt} &= -I/C \\ \frac{dI}{dt} &= -(IR - V)/L = V/L - IR/L\end{aligned}$$

or in matrix vector form

$$\mathbf{x}' = A\mathbf{x}$$

where $\mathbf{x} = (V, I)^T$ and

$$A = \begin{bmatrix} 0 & -1/C \\ 1/L & -R/L \end{bmatrix}$$

The eigenvalues λ of A satisfy

$$\det(A - \lambda I) = \lambda^2 + \frac{R}{L}\lambda + \frac{1}{LC} = 0$$

with solutions

$$\lambda = \frac{1}{2L} \left(-R \pm \sqrt{R^2 - 4L/C} \right).$$

Considering this expression carefully you can see that if the λ values are real they are negative. If they are complex, the real part is $-R/(2L)$. Thus solutions always decay exponentially. This is to be expected in a circuit with no external power. For the circuit to have oscillations, the eigenvalues must be complex, so

$$R^2 - 4L/C < 0$$

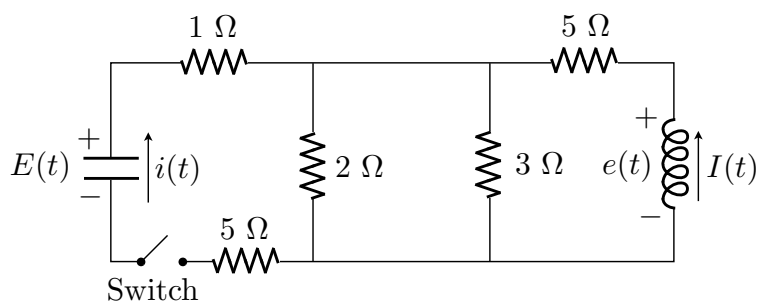


Figure 6.4: Circuit diagram for Example 6.16.

which is commonly rewritten as

$$\frac{R}{2} \sqrt{\frac{C}{L}} < 1.$$

Example 6.16 Consider the circuit shown in Figure 6.4. The capacitor has capacitance $C = 1$ Farad and the inductor inductance $L = 1$ Henry. The capacitor is charged up to $12V$ and then the switch is closed (so $E(0) = 12$ and $I(0) = 0$).

- (a) Derive a system of differential equations for $I(t)$ and $E(t)$.
- (b) We expect that $E(t)$ and $I(t) \rightarrow 0$ as $t \rightarrow \infty$ (no external power). Will there be oscillations in E and I in time?

Remember that if $i(t)$ is the current upward through the capacitor then

$$\frac{dE}{dt} = -\frac{i}{C} = -i \quad (6.11)$$

since $C = 1$. If $e(t)$ is the voltage across the inductor as shown then

$$\frac{dI}{dt} = -\frac{e}{L} = -e \quad (6.12)$$

since $L = 1$. It is still necessary to work out $i(t)$ and $e(t)$ in terms of $E(t)$ and $I(t)$. This is the fundamental problem for the circuit. We solved this problem in Example 3.13 in Section 3.6 and considered it from a different perspective in the additional notes to Chapter 4. We found that

$$\begin{aligned} i &= \frac{5}{36}E - \frac{1}{6}I \\ e &= \frac{1}{6}E + 6I \end{aligned}$$

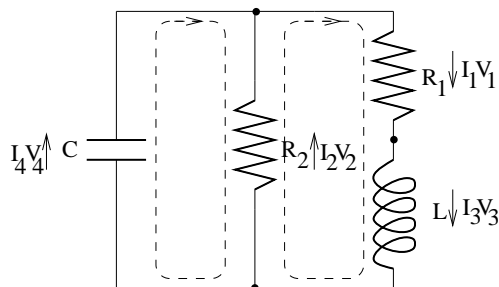


Figure 6.5: Circuit diagram.

Inserting these into (6.11) and (6.12) gives the desired system for E and I :

$$\begin{aligned}\frac{dE}{dt} &= -\frac{5}{36}E + \frac{1}{6}I \\ \frac{dI}{dt} &= -\frac{1}{6}E - 6I\end{aligned}$$

or in vector form $\mathbf{x}' = A\mathbf{x}$ where

$$\mathbf{x} = \begin{bmatrix} E \\ I \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} -5/36 & 1/6 \\ -1/6 & -6 \end{bmatrix}$$

An eigenanalysis of A gives $\lambda_1 \approx -0.1436$ and $\lambda_2 \approx -5.9953$. Since these are not complex, the circuit does not exhibit oscillations.

6.4.3 Alternate description of LCR circuits

LCR circuits were presented differently in a previous version of the notes. That approach is reproduced below starting in the next paragraph. It follows the discussion in Section 3.5.4.

We now return to the circuit that we discussed previously, shown in Figure 6.5. Recall that we chose as basic variables I_3 and V_4 and solved for all the other variables in terms of these. The result was

$$\begin{aligned}I_1 &= I_3 \\ I_2 &= \frac{1}{R_2}V_4 \\ I_4 &= I_3 - \frac{1}{R_2}V_4 \\ V_1 &= R_1I_3 \\ V_2 &= V_4 \\ V_3 &= -R_1I_3 - V_4\end{aligned}$$

Now we can complete the job and determine I_3 and V_4 . We have to take into account now that the currents and voltages are functions of time. The relations

between currents and voltages across capacitors and inductors involves the time derivatives.

If I and V are the current flowing through and the voltage across a capacitor with capacitance C , then

$$\frac{dV}{dt} = \frac{1}{C}I$$

If I and V are the current flowing through and the voltage across an inductor with inductance L , then

$$\frac{dI}{dt} = \frac{1}{L}V$$

Notice that we have chosen as basic the variables that get differentiated.

Using these relations for I_3 and V_4 yields

$$\begin{aligned}\frac{dI_3}{dt} &= \frac{1}{L}V_3 \\ \frac{dV_4}{dt} &= \frac{1}{C}I_4\end{aligned}$$

Now we re-express everything in terms of I_3 and V_4 using the equations we derived previously.

$$\begin{aligned}\frac{dI_3}{dt} &= \frac{1}{L}(-R_1I_3 - V_4) = \frac{-R_1}{L}I_3 - \frac{1}{L}V_4 \\ \frac{dV_4}{dt} &= \frac{1}{C}(I_3 - \frac{1}{R_2}V_4) = \frac{1}{C}I_3 - \frac{1}{R_2C}V_4\end{aligned}$$

This can be written as

$$\begin{bmatrix} I_3 \\ V_4 \end{bmatrix}' = \begin{bmatrix} \frac{-R_1}{L} & -\frac{1}{L} \\ \frac{1}{C} & -\frac{1}{R_2C} \end{bmatrix} \begin{bmatrix} I_3 \\ V_4 \end{bmatrix}$$

Lets try to determine for what values of R_1 , L , C and R_2 the circuit exhibits oscillations. Recall that the solution will involve sines and cosines whenever the matrix has complex eigenvalues.

The polynomial $\det(A - \lambda I) = \lambda^2 + b\lambda + c$, where

$$b = \frac{R_1}{L} + \frac{1}{R_2C}$$

and

$$c = \frac{R_1}{R_2LC} + \frac{1}{LC}.$$

The eigenvalues are the roots of this polynomial, given by $(-b \pm \sqrt{b^2 - 4c})/2$. These will be complex if $b^2 < 4c$, i.e., if

$$\left(\frac{R_1}{L} + \frac{1}{R_2C}\right)^2 < 4\left(\frac{R_1}{R_2LC} + \frac{1}{LC}\right)$$

Notice that this can be achieved by decreasing R_1 and increasing R_2

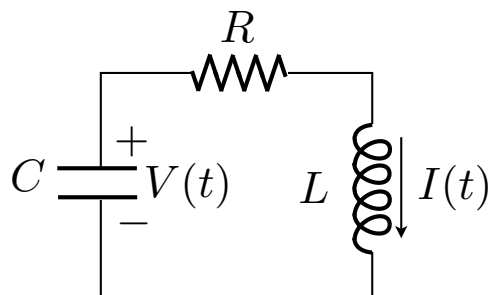


Figure 6.6: Circuit diagram for Problem 6.23.

6.4.4 Problems

Problem 6.22 In the circuit in Figure 6.5, suppose that $R_1 = R_2 = 1$ ohm, $C = 1$ farad and $L = 1$ henry. If the initial current across the inductor is $I_3(0) = 1$ ampere and initial voltage across the capacitor is $V_4(0) = 1$ volt, find $I_3(t)$ and $V_4(t)$ for all later times. What is $V_1(t)$?

Problem 6.23 Consider the circuit with diagram in Figure 6.6 Write down the system of equations satisfied by I and V . For what values of L , C and R does the circuit exhibit oscillations? Suppose that $R = 1$ ohm, $C = 1$ farad and $L = 1$ henry. If the initial current across the inductor is $I(0) = 1$ ampere and initial voltage across the capacitor is $V(0) = 1$ volt, find $I(t)$ and $V(t)$ for all later times.

6.5 Additional Topics

6.5.1 Diagonalization

Diagonal matrices (that is, matrices that have zero entries except on the diagonal) are extremely easy to work with. For a start, the eigenvalues of a diagonal matrix are exactly the diagonal entries. If

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

then $\det(D - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda)$ which is zero precisely when λ equals one of $\lambda_1, \lambda_2, \dots, \lambda_n$. The corresponding eigenvectors are just the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$.

It is also easy to compute powers of a diagonal matrix. We simply obtain

$$D^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}$$

This formula makes it easy to compute the matrix exponential of D . Recall that the matrix e^{tD} is defined to be the matrix power series

$$e^{tD} = I + tD + \frac{t^2}{2}D^2 + \frac{t^3}{3!}D^3 + \cdots$$

Using the formula above we find that

$$\begin{aligned} e^{tD} &= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} + \begin{bmatrix} t\lambda_1 & 0 & \cdots & 0 \\ 0 & t\lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & t\lambda_n \end{bmatrix} + \begin{bmatrix} \frac{t^2\lambda_1^2}{2} & 0 & \cdots & 0 \\ 0 & \frac{t^2\lambda_2^2}{2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{t^2\lambda_n^2}{2} \end{bmatrix} + \cdots \\ &= \begin{bmatrix} 1 + t\lambda_1 + \frac{t^2\lambda_1^2}{2} + \cdots & 0 & \cdots & 0 \\ 0 & 1 + t\lambda_2 + \frac{t^2\lambda_2^2}{2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 + t\lambda_n + \frac{t^2\lambda_n^2}{2} + \cdots \end{bmatrix} \\ &= \begin{bmatrix} e^{t\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{t\lambda_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & e^{t\lambda_n} \end{bmatrix} \end{aligned}$$

Things are not quite so simple for an arbitrary $n \times n$ matrix A . However, if A has a basis of eigenvectors then it turns out that there exists an invertible matrix B such that $AB = DB$, where D is the diagonal matrix whose diagonal elements are the eigenvalues of A . Multiplying by B^{-1} from either the left or right gives

$$A = BDB^{-1}, \quad D = B^{-1}AB$$

In fact, B is simply the matrix whose columns are the eigenvectors of A . In other words, if $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ are the eigenvectors for A then $B = [\mathbf{x}_1 | \mathbf{x}_2 | \cdots | \mathbf{x}_n]$. To see this notice that

$$\begin{aligned} AB &= A[\mathbf{x}_1 | \mathbf{x}_2 | \cdots | \mathbf{x}_n] \\ &= [A\mathbf{x}_1 | A\mathbf{x}_2 | \cdots | A\mathbf{x}_n] \\ &= [\lambda_1\mathbf{x}_1 | \lambda_2\mathbf{x}_2 | \cdots | \lambda_n\mathbf{x}_n] \end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \\
&= BD
\end{aligned}$$

The assumption that A has a basis of eigenvectors implies that the matrix B is invertible.

Using the representation $A = BDB^{-1}$ it is easy to calculate powers of A . We have

$$A^2 = BDB^{-1}BDB^{-1} = BDIDB^{-1} = BD^2B^{-1}$$

and similarly

$$A^k = BD^k B^{-1}$$

Therefore we can now also sum the power series for the exponential and obtain

$$e^{tA} = Be^{tD}B^{-1}$$

6.5.2 Computing high powers of a matrix

Recall that when we were discussing the random walk problem, we ended up with the problem of computing the limit for large n of $P^n \mathbf{x}_0$ where P is the matrix of transition probabilities.

We can now solve this problem using diagonalization. Lets do an example. Suppose that

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

We wish to compute P^n for large n .

We begin by diagonalizing P . This involves finding the eigenvalues and eigenvectors. I won't give the details of this computation. The results are $\lambda_1 = 1$, $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 8/3 \end{bmatrix}$ and $\lambda_2 = 1/12$, $\mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. So

$$P = \begin{bmatrix} 1 & -1 \\ \frac{8}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{12} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -1 \\ \frac{8}{3} & 1 \end{bmatrix}^{-1}$$

and

$$P^n = \begin{bmatrix} 1 & -1 \\ \frac{8}{3} & 1 \end{bmatrix} \begin{bmatrix} 1^n & 0 \\ 0 & (\frac{1}{12})^n \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -1 \\ \frac{8}{3} & 1 \end{bmatrix}^{-1}$$

But $1^n = 1$ for all n and $(\frac{1}{12})^n \rightarrow 0$ as $n \rightarrow \infty$, since $\frac{1}{12} < 1$. So

$$\begin{aligned}
\lim_{n \rightarrow \infty} P^n &= \begin{bmatrix} 1 & -1 \\ \frac{8}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -1 \\ \frac{8}{3} & 1 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 1 & -1 \\ \frac{8}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{3}{11} & \frac{3}{11} \\ \frac{3}{11} & \frac{3}{11} \end{bmatrix} \\
&= \begin{bmatrix} \frac{3}{11} & \frac{3}{11} \\ \frac{3}{11} & \frac{3}{11} \end{bmatrix}
\end{aligned}$$

6.5.3 Another formula for the determinant

If A has a basis of eigenvectors, then we can get another formula for the determinant. Using the multiplicative property of the determinant, we have

$$\det(A) = \det(BDB^{-1}) = \det(B) \det(D) \det(B)^{-1} = \det(D).$$

But $\det(D)$ is just the product of the diagonal elements, i.e., the eigenvalues. Thus the determinant of A is the product of its eigenvalues:

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Actually, this is true even if A doesn't have a basis of eigenvectors and isn't diagonalizable.

6.5.4 The matrix exponential and differential equations

The matrix exponential e^{tA} can be used to solve the differential equation

$$\mathbf{y}'(t) = A\mathbf{y}(t)$$

with initial condition

$$\mathbf{y}(0) = \mathbf{x}_0$$

To see this notice that e^{tA} satisfies the differential equation $\frac{d}{dt}e^{tA} = Ae^{tA}$. This follows from the power series representation

$$e^{tA} = I + tA + \frac{t^2}{2}A^2 + \frac{t^3}{3!}A^3 + \cdots$$

since

$$\begin{aligned} \frac{d}{dt}e^{tA} &= A + \frac{2t}{2}A^2 + \frac{3t^2}{3!}A^3 + \cdots \\ &= A + tA^2 + \frac{t^2}{2!}A^3 + \cdots \\ &= A(I + tA + \frac{t^2}{2}A^2 + \cdots) \\ &= Ae^{tA} \end{aligned}$$

Also

$$e^{0A} = I$$

These two facts imply that $\mathbf{y}(t) = e^{tA}\mathbf{x}_0$ is the solution to our differential equation and initial condition, since $\mathbf{y}'(t) = \frac{d}{dt}e^{tA}\mathbf{x}_0 = Ae^{tA}\mathbf{x}_0 = A\mathbf{y}(t)$ and $\mathbf{y}(0) = e^{0A}\mathbf{x}_0 = I\mathbf{x}_0 = \mathbf{x}_0$.

The matrix exponential is a nice theoretical construction. However, to actually compute the matrix exponential using diagonalization involves just the same ingredients—computing the eigenvalues and vectors—as our original solution. In fact it is more work.

However, there is one situation where the matrix exponential gives us something new. This is the situation where A does not have a basis of eigenvectors. The power series definition of the matrix exponential still makes sense, and can compute it in certain special cases. Consider the matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. This matrix does not have a basis of eigenvectors. So it cannot be diagonalized. However, in a homework problem, you showed that $e^{tA} = \begin{bmatrix} e^t & te^t \\ 0 & e^t \end{bmatrix}$. Thus the solution to

$$\mathbf{y}'(t) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{y}(t)$$

with initial condition

$$\mathbf{y}(0) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

is

$$\mathbf{y}(t) = e^{tA} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} e^t & te^t \\ 0 & e^t \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2e^t + te^t \\ e^t \end{bmatrix}$$

Notice that this solution involves a power of t in addition to exponentials.

6.5.5 Converting higher order equations into first order systems

So far we have only considered first order differential equations. In other words, the equations have only involved first derivatives $y'(t)$ and not higher derivatives like $y''(t)$. However higher order equations, especially second order equations, occur often in practical problems. In this section we will show that a higher order linear differential equation can be converted into an equivalent first order system.

Suppose we want to solve the equation

$$y''(t) + ay'(t) + by(t) = 0$$

with initial conditions

$$\begin{aligned} y(0) &= y_0 \\ y'(0) &= y'_0 \end{aligned}$$

Define the new functions $z_1(t)$ and $z_2(t)$ by

$$\begin{aligned} z_1(t) &= y(t) \\ z_2(t) &= y'(t) \end{aligned}$$

Then

$$\begin{aligned} z'_1(t) &= y'(t) &= z_2(t) \\ z'_2(t) &= y''(t) &= -ay'(t) - by(t) &= -az_2(t) - bz_1(t) \end{aligned}$$

and

$$\begin{aligned} z_1(0) &= y_0 \\ z_2(0) &= y'_0 \end{aligned}$$

In other words the vector $\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}$ satisfies the equation

$$\frac{d}{dt} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}$$

with initial condition

$$\begin{bmatrix} z_1(0) \\ z_2(0) \end{bmatrix} = \begin{bmatrix} y_0 \\ y'_0 \end{bmatrix}.$$

Example 6.17 Suppose we want to solve the second order equation

$$y'' + 4y' + y = 0$$

with initial conditions

$$y(0) = 1, y'(0) = 0$$

If we let $z_1(t) = y(t)$ and $z_2(t) = y'(t)$ then

$$\frac{d}{dt} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}$$

with initial condition

$$\begin{bmatrix} z_1(0) \\ z_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

To solve this we first find the eigenvalues and eigenvectors. They are $\lambda_1 = -2 + \sqrt{3}$, $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -2 + \sqrt{3} \end{bmatrix}$ and $\lambda_2 = -2 - \sqrt{3}$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -2 - \sqrt{3} \end{bmatrix}$. So the general solution is

$$c_1 e^{(-2+\sqrt{3})t} \begin{bmatrix} 1 \\ -2 + \sqrt{3} \end{bmatrix} + c_2 e^{(-2-\sqrt{3})t} \begin{bmatrix} 1 \\ -2 - \sqrt{3} \end{bmatrix}$$

To satisfy the initial condition, we need

$$c_1 \begin{bmatrix} 1 \\ -2 + \sqrt{3} \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -2 - \sqrt{3} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The solution is

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \sqrt{3}/3 + 1/2 \\ -\sqrt{3}/3 + 1/2 \end{bmatrix}$$

Thus

$$\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = (\sqrt{3}/3 + 1/2)e^{(-2+\sqrt{3})t} \begin{bmatrix} 1 \\ -2 + \sqrt{3} \end{bmatrix} + (-\sqrt{3}/3 + 1/2)e^{(-2-\sqrt{3})t} \begin{bmatrix} 1 \\ -2 - \sqrt{3} \end{bmatrix}$$

and so

$$y(t) = z_1(t) = (\sqrt{3}/3 + 1/2)e^{(-2+\sqrt{3})t} + (-\sqrt{3}/3 + 1/2)e^{(-2-\sqrt{3})t}$$

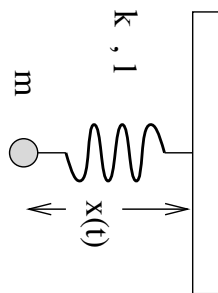


Figure 6.7: A simple mass-spring system.

Actually, to solve the equation

$$y''(t) + ay'(t) + by(t) = 0$$

it is not really necessary to turn it into a first order system. We can simply try to find solutions of the form $y(t) = e^{\lambda t}$. If we plug this into the equation we get $(\lambda^2 + a\lambda + b)e^{\lambda t}$ which is zero if λ is a root of $\lambda^2 + a\lambda + b$. This polynomial has two roots, which yields two solutions.

Still, the procedure of turning a higher order equation into a first order system is important. This is because on a computer it is much easier to solve a first order system than a high order equation. If the coefficients a and b are functions of t , then exact solutions (like exponentials) usually can't be found. However, one can still turn the equation into a first order system $\mathbf{y}'(t) = A(t)\mathbf{y}(t)$ where the matrix now depends on t and solve this on a computer.

6.5.6 Springs and weights

To begin, let's consider the situation where we have a single weight hanging on a spring shown in Figure 6.7. We want to determine how the weight moves in time. To do this we calculate the forces acting on the weight and use Newton's law of motion $F = ma$.

One force acting on the weight are the force of gravity. This acts in the positive x direction (i.e., downward) and has magnitude mg . The other force is due to the spring. It's magnitude is $k(x - l)$ in the negative x direction. The acceleration is the second derivative $x''(t)$. Thus the total force is $F = mg - k(x(t) - l)$ and $ma = mx''(t)$ Newton's law reads

$$mx''(t) = mg - k(x - l) = -kx + mg + lk$$

This is not quite in the form we can handle, due to the term $mg + lk$ on the right. What we must do is first find the equilibrium solution. In a previous lecture we found the equilibrium position by minimizing the potential energy. There is another, equivalent, way. That is to find the value of x for which the total force is zero. In other words

$$-kx_{\text{eq}} + mg + lk = 0$$

or

$$x_{\text{eq}} = (mg + lk)/k$$

Notice that the total force can be written

$$-kx + mg + lk = -k(x - x_{\text{eq}})$$

Now let $y(t) = x(t) - x_{\text{eq}}$ be the displacement from the equilibrium point. Notice that $y'(t) = x'(t)$ and $y''(t) = x''(t)$, since x_{eq} is a constant. So the equation for $y(t)$ is

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

or

$$y''(t) + \frac{k}{m}y(t) = 0$$

We could turn this into a first order system. However, it is easier to try solutions of the form $e^{\lambda t}$. Substituting this into the equation yields

$$(\lambda^2 + k/m)e^{\lambda t} = 0$$

so we require that $\lambda^2 + k/m = 0$, or $\lambda = \pm i\sqrt{k/m}$. Thus, solutions are $e^{i\sqrt{k/m}t}$ and $e^{-i\sqrt{k/m}t}$. To obtain real solutions, we can take the real and imaginary parts. This gives as solutions $\sin(\sqrt{k/m}t)$ and $\cos(\sqrt{k/m}t)$, and the general solution is

$$c_1 \sin(\sqrt{k/m}t) + c_2 \cos(\sqrt{k/m}t)$$

We can make the equation a little more interesting by adding friction. A frictional force is proportional to the velocity and acts in the direction opposite to the motion. Thus the equation with friction reads

$$y''(t) + \beta y'(t) + \frac{k}{m}y(t) = 0$$

This can be solved by turning it into a first order system, or directly, using trial solution of the form $e^{\lambda t}$ as above.

Now we turn the problem with several weights and springs shown in Figure 6.8. In this problem matrices play an essential role.

We begin by computing the forces acting on each weight. Let us start with the first weight. Gravity is pulling down, and the springs above and below are pushing or pulling with a force proportional to their extensions. Thus the total force on the first weight is $m_1g - k_1(x_1 - l_1) + k_2(x_2 - x_1 - l_2)$. To get the signs right on the spring forces, think of what happens when one of the x_i 's gets large. For example, when x_1 gets large, the first spring stretches and pulls up, so the sign of the force should be negative for large x_1 . So Newton's equation for the first weight is

$$m_1x_1''(t) = m_1g - k_1(x_1 - l_1) + k_2(x_2 - x_1 - l_2) = -(k_1 + k_2)x_1 + k_2x_2 + m_1g + k_1l_1 - k_2l_2$$

or

$$x_1''(t) = -\frac{k_1 + k_2}{m_1}x_1 + \frac{k_2}{m_1}x_2 + g + \frac{k_1l_1 - k_2l_2}{m_1}$$

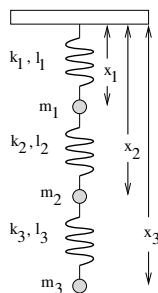


Figure 6.8: A more complicated mass-spring system.

Similarly the equations for the second and third weights are

$$\begin{aligned}x_2''(t) &= \frac{k_2}{m_2}x_1 - \frac{k_2 + k_3}{m_2}x_2 + \frac{k_3}{m_2}x_3 + g + \frac{k_2l_2 - k_3l_3}{m_2} \\x_3''(t) &= \frac{k_3}{m_3}x_2 - \frac{k_3}{m_3}x_3 + g + \frac{k_3l_3}{m_3}\end{aligned}$$

Thus can be written as a second order matrix equation

$$\mathbf{x}''(t) = K\mathbf{x}(t) + \mathbf{b}$$

where

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix},$$

$$K = \begin{bmatrix} -\frac{k_1+k_2}{m_1} & \frac{k_2}{m_1} & 0 \\ \frac{k_2}{m_2} & -\frac{k_2+k_3}{m_2} & \frac{k_3}{m_2} \\ 0 & \frac{k_3}{m_3} & -\frac{k_3}{m_3} \end{bmatrix}$$

and

$$\mathbf{b} = \begin{bmatrix} g + \frac{k_1l_1 - k_2l_2}{m_1} \\ g + \frac{k_2l_2 - k_3l_3}{m_2} \\ g + \frac{k_3l_3}{m_3} \end{bmatrix}.$$

With this notation, the equilibrium solution is the value of \mathbf{x} that makes all the forces zero. That is,

$$K\mathbf{x}_{\text{eq}} + \mathbf{b} = \mathbf{0}$$

or,

$$\mathbf{x}_{\text{eq}} = -K^{-1}\mathbf{b}$$

As in the case of one weight the force side of the equation can now be written as

$$K\mathbf{x} + \mathbf{b} = K(\mathbf{x} + K^{-1}\mathbf{b}) = K(\mathbf{x} - \mathbf{x}_{\text{eq}})$$

so if we define

$$\mathbf{y}(t) = \mathbf{x}(t) - \mathbf{x}_{\text{eq}},$$

the equation for $\mathbf{y}(t)$ is

$$\mathbf{y}''(t) = K\mathbf{y}(t)$$

To solve this second order 3×3 system, we could turn it to a first order 6×6 system. However, just as in the case of a single higher order equation we can proceed directly. We try trial solutions of the form $e^{\kappa t}\mathbf{y}$. Substituting this into the equation, we see that the equation is satisfied if

$$\kappa^2\mathbf{y} = K\mathbf{y}$$

in other words, κ^2 is an eigenvalue of K with eigenvector \mathbf{y} , or κ is one of the two square roots of an eigenvalue.

So, if K has eigenvalues λ_1 , λ_2 and λ_3 with eigenvectors \mathbf{y}_1 , \mathbf{y}_2 and \mathbf{y}_3 , then six solutions of the equation are given by $e^{\sqrt{\lambda_1}t}\mathbf{y}_1$, $e^{-\sqrt{\lambda_1}t}\mathbf{y}_1$, $e^{\sqrt{\lambda_2}t}\mathbf{y}_2$, $e^{-\sqrt{\lambda_2}t}\mathbf{y}_2$, $e^{\sqrt{\lambda_3}t}\mathbf{y}_3$ and $e^{-\sqrt{\lambda_3}t}\mathbf{y}_3$. If some of the λ_i 's are negative, then these solutions are complex exponentials, and we may take their real and imaginary parts to get real solutions. The general solution is a linear combination of these, and the coefficients in the linear combination may be chosen to satisfy an initial condition.

To make this clear we will do an example. Suppose that all the masses m_i , lengths l_i and spring constants k_i are equal to 1. Then

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

Suppose that the initial position of the weights is $x_1 = 30$, $x_2 = 60$ and $x_3 = 70$, and that the initial velocities are $x'_1 = 1$ and $x'_2 = x'_3 = 0$. We will determine the positions of the weights for all subsequent times. The numbers in this problem don't turn out particularly nicely, so I'll just give them to 3 significant figures.

The first step is to find the eigenvalues and eigenvectors of K . They are given by

$$\lambda_1 = -0.198 \quad \lambda_2 = -1.55 \quad \lambda_3 = -3.25$$

$$\mathbf{x}_1 = \begin{bmatrix} 0.445 \\ 0.802 \\ 1.00 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} -1.25 \\ -0.555 \\ 1.00 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 1.80 \\ -2.25 \\ 1.00 \end{bmatrix}$$

Let $\mu_1 = \sqrt{0.198} = 0.445$, $\mu_2 = \sqrt{1.55} = 1.25$ and $\mu_3 = \sqrt{3.25} = 1.80$. Then if $\mathbf{y}(t) = \mathbf{x}(t) - \mathbf{x}_{\text{eq}}$, then general solution for $\mathbf{y}(t)$ is

$$\mathbf{y}(t) = (c_1 e^{i\mu_1 t} + d_1 e^{-i\mu_1 t})\mathbf{x}_1 + (c_2 e^{i\mu_2 t} + d_2 e^{-i\mu_2 t})\mathbf{x}_2 + (c_3 e^{i\mu_3 t} + d_3 e^{-i\mu_3 t})\mathbf{x}_3$$

where $c_1, d_1, c_2, d_2, c_3, d_3$ are arbitrary constants. Taking real and imaginary parts, we can also write the general solution as

$$\mathbf{y}(t) = (a_1 \cos(\mu_1 t) + b_1 \sin(\mu_1 t))\mathbf{x}_1 + (a_2 \cos(\mu_2 t) + b_2 \sin(\mu_2 t))\mathbf{x}_2 + (a_3 \cos(\mu_3 t) + b_3 \sin(\mu_3 t))\mathbf{x}_3$$

where $a_1, b_1, a_2, b_2, a_3, b_3$ are arbitrary constants. Notice that we can find the general solution for $\mathbf{y}(t) = \mathbf{x}(t) - \mathbf{x}_{\text{eq}}$ without knowing \mathbf{x}_{eq} . However, since the initial conditions were given in terms of \mathbf{x} and not \mathbf{y} , we now have to find \mathbf{x}_{eq} to be able to convert initial conditions for \mathbf{x} to initial conditions for \mathbf{y} . If we work in units where $g = 10$ then

$$\mathbf{b} = \begin{bmatrix} 10 \\ 10 \\ 11 \end{bmatrix}$$

and

$$\mathbf{x}_{\text{eq}} = -K^{-1}\mathbf{b} = \begin{bmatrix} 31 \\ 52 \\ 63 \end{bmatrix}$$

so

$$\mathbf{y}(0) = \mathbf{x}(0) - \mathbf{x}_{\text{eq}} = \begin{bmatrix} 30 \\ 60 \\ 70 \end{bmatrix} - \begin{bmatrix} 31 \\ 52 \\ 63 \end{bmatrix} = \begin{bmatrix} -1 \\ 8 \\ 7 \end{bmatrix}$$

Also

$$\mathbf{y}'(0) = \mathbf{x}'(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

So to satisfy the first initial condition, since $\cos(0) = 1$ and $\sin(0) = 0$, we need that

$$\mathbf{y}(0) = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + a_3\mathbf{x}_3 = \begin{bmatrix} -1 \\ 8 \\ 7 \end{bmatrix}.$$

Explicitly, we need to solve

$$[\mathbf{x}_1 | \mathbf{x}_2 | \mathbf{x}_3] \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 8 \\ 7 \end{bmatrix},$$

or

$$\begin{bmatrix} 0.445 & -1.25 & 1.80 \\ 0.802 & -0.555 & -2.25 \\ 1.00 & 1.00 & 1.00 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 8 \\ 7 \end{bmatrix}$$

This is not a pretty sight, but I can punch the numbers into my computer and find that

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 7.04 \\ 1.33 \\ -1.37 \end{bmatrix}$$

To satisfy the second initial condition, we differentiate the expression for the general solution of $\mathbf{y}(t)$ and set $t = 0$. This gives

$$\mu_1 b_1 \mathbf{x}_1 + \mu_2 b_2 \mathbf{x}_2 + \mu_3 b_3 \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Solving this numerically gives

$$\begin{bmatrix} \mu_1 b_1 \\ \mu_2 b_2 \\ \mu_3 b_3 \end{bmatrix} = \begin{bmatrix} 0.242 \\ -0.435 \\ 0.194 \end{bmatrix}$$

Finally, we divide by the μ_i 's to give

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0.543 \\ -0.348 \\ 1.80 \end{bmatrix}$$

Now we have completely determined all the constants, so the solution is complete.

6.5.7 Problems

Problem 6.24 Consider the second order equation

$$y'' - 5y' + 6y = 0$$

with initial conditions $y(0) = 1$, $y'(0) = 0$. Solve this by turning it into a 2×2 system. Then solve it directly by using trial solutions $e^{\lambda t}$.

Problem 6.25 Consider the second order equation

$$y'' + y' + y = 0$$

with initial conditions $y(0) = 1$, $y'(0) = 0$. Solve this by turning it into a 2×2 system. Then solve it directly by using trial solutions $e^{\lambda t}$.

Problem 6.26 How can you turn a third order equation

$$y''' + ay'' + by' + cy = 0$$

into an equivalent 3×3 system of equations?

Problem 6.27 Suppose K is a 3×3 matrix with eigenvalues and eigenvectors given by

$$\lambda_1 = -1 \quad \lambda_2 = -4 \quad \lambda_3 = -9$$

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Find the solution of

$$\mathbf{y}''(t) = K\mathbf{y}(t)$$

satisfying

$$\mathbf{y}(0) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\mathbf{y}'(0) = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Problem 6.28 Consider a system of two hanging weights and springs. Suppose that all the masses, spring constants and spring lengths are equal to one, and that $g = 10$. Find the positions $x_1(t)$ and $x_2(t)$ for all times if $x_1(0) = 20$, $x_2(0) = 30$, $x_1'(0) = 1$, $x_2'(0) = 1$.

6.6 Solutions to Chapter Problems

Solution 6.1 We compute $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, so $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue 2. Also $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, so $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector with eigenvalue 0.

Solution 6.2 If P projects onto some line, then a vector \mathbf{x} lying on that line doesn't get changed by P so $P\mathbf{x} = \mathbf{x}$ and \mathbf{x} is an eigenvector with eigenvalue 1. On the other hand, if \mathbf{x} is perpendicular to the line, then $P\mathbf{x} = \mathbf{0} = 0\mathbf{x}$ so \mathbf{x} is an eigenvector with eigenvalue 0.

Solution 6.3

- a) $\det(A - \lambda I) = \lambda^2 - 9 = (\lambda - 3)(\lambda + 3)$ so the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = -3$. To find the eigenvector for $\lambda_1 = 3$ we must solve the homogeneous equation with matrix $\begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} -3 & 3 \\ 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. To find the eigenvector for $\lambda_2 = -3$ we must solve the homogeneous equation with matrix $\begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} 3 & 3 \\ 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.
- b) $\det(A - \lambda I) = \lambda^2 - 8\lambda + 12 = (\lambda - 2)(\lambda - 6)$ so the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 6$. To find the eigenvector for $\lambda_1 = 2$ we must solve the homogeneous equation with matrix $\begin{bmatrix} -4 & -8 \\ 4 & 8 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} -4 & -8 \\ 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. To find the eigenvector for $\lambda_2 = 6$ we must solve the homogeneous equation with matrix $\begin{bmatrix} -8 & -8 \\ 4 & 4 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} -8 & -8 \\ 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.
- c) $\det(A - \lambda I) = \lambda^2 + 7\lambda + 6 = (\lambda + 6)(\lambda + 1)$ so the eigenvalues are $\lambda_1 = -6$ and $\lambda_2 = -1$. To find the eigenvector for $\lambda_1 = -6$ we must solve the

homogeneous equation with matrix $\begin{bmatrix} 35 & -10 \\ 105 & -30 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} 35 & -10 \\ 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 2 \\ 7 \end{bmatrix}$. To find the eigenvector for $\lambda_2 = -1$ we must solve the homogeneous equation with matrix $\begin{bmatrix} 30 & -10 \\ 105 & -35 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} 30 & -10 \\ 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

d) $\det(A - \lambda I) = \lambda^2 - 3\lambda - 10 = (\lambda + 2)(\lambda - 5)$ so the eigenvalues are $\lambda_1 = -2$ and $\lambda_2 = 5$. To find the eigenvector for $\lambda_1 = -2$ we must solve the homogeneous equation with matrix $\begin{bmatrix} -7 & -14 \\ 7 & 14 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} -7 & -14 \\ 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. To find the eigenvector for $\lambda_2 = 5$ we must solve the homogeneous equation with matrix $\begin{bmatrix} -14 & -14 \\ 7 & 7 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} -14 & -14 \\ 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Solution 6.4

$$\begin{aligned} \det \begin{bmatrix} 2 - \lambda & 3 \\ 2 & 1 - \lambda \end{bmatrix} &= (2 - \lambda)(1 - \lambda) - 6 = 2 - 3\lambda + \lambda^2 - 6 \\ &= \lambda^2 - 3\lambda - 4 = (\lambda - 4)(\lambda + 1) = 0 \\ &\Rightarrow \lambda = -1, 4. \end{aligned}$$

The eigenvector for $\lambda = -1$:

$$\begin{bmatrix} 3 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \vec{0} \quad \Rightarrow \quad \begin{aligned} 3k_1 + 3k_2 &= 0 \\ 2k_1 + 2k_2 &= 0 \end{aligned} \quad \Rightarrow \quad k_1 = -k_2$$

If we set $k_2 = 1$ for convenience, we have that the first eigenpair is

$$\lambda_1 = -1, \quad \vec{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Now, the eigenvector for $\lambda = 4$:

$$\begin{bmatrix} -2 & 3 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \vec{0} \quad \Rightarrow \quad \begin{aligned} -2k_1 + 3k_2 &= 0 \\ 2k_1 - 3k_2 &= 0 \end{aligned} \quad \Rightarrow \quad k_1 = \frac{3}{2}k_2$$

If we set $k_2 = 1$ for convenience, we have that the second eigenpair is

$$\lambda_2 = 4, \quad \vec{v}_2 = \begin{bmatrix} 3/2 \\ 1 \end{bmatrix}.$$

Solution 6.5

a) $\det(A - \lambda I) = -\lambda^3 + 2\lambda^2 + \lambda - 2 = -(\lambda - 1)(\lambda - 2)(\lambda + 1)$ so the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2$ and $\lambda_3 = -1$. To find the eigenvector for $\lambda_1 = 1$

we must solve the homogeneous equation with matrix $\begin{bmatrix} -1 & -1 & 1 \\ 1 & -1 & 2 \\ 2 & 0 & 1 \end{bmatrix}$. The

matrix reduces to $\begin{bmatrix} -1 & -1 & 1 \\ 0 & -2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$. To

find the eigenvector for $\lambda_2 = 2$ we must solve the homogeneous equation with matrix $\begin{bmatrix} -2 & -1 & 1 \\ 1 & -2 & 2 \\ 2 & 0 & 0 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} -2 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and the

eigenvector is $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$. To find the eigenvector for $\lambda_3 = -1$ we must solve

the homogeneous equation with matrix $\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 2 \\ 2 & 0 & 3 \end{bmatrix}$. The matrix reduces

to $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_3 = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$.

b) $\det(A - \lambda I) = -\lambda^3 + 2\lambda^2 + 3\lambda - 6 = -(\lambda - 2)(\lambda - \sqrt{3})(\lambda + \sqrt{3})$ so the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = \sqrt{3}$ and $\lambda_3 = -\sqrt{3}$. To find the eigenvector for $\lambda_1 = 2$ we must solve the homogeneous equation with matrix

$\begin{bmatrix} -1 & 1 & 1 \\ 1 & -2 & -2 \\ 1 & -1 & -1 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} 0 & -1 & -1 \\ 0 & 0 & 0 \\ -1 & 1 & 1 \end{bmatrix}$ and the eigen-

vector is $\mathbf{x}_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$. To find the eigenvector for $\lambda_2 = \sqrt{3}$ we must solve

the homogeneous equation with matrix $\begin{bmatrix} 1 - \sqrt{3} & 1 & 1 \\ 1 & -\sqrt{3} & -2 \\ 1 & -1 & 1 - \sqrt{3} \end{bmatrix}$. The

matrix reduces to $\begin{bmatrix} 1 - \sqrt{3} & 1 & 1 \\ 0 & 2 - \sqrt{3} & -3 + 2\sqrt{3} \\ 0 & 0 & 0 \end{bmatrix}$ and the eigenvector is

$\mathbf{x}_2 = \begin{bmatrix} -1 \\ -\sqrt{3} \\ 1 \end{bmatrix}$. To find the eigenvector for $\lambda_3 = -\sqrt{3}$ we must solve the ho-

mogenuous equation with matrix $\begin{bmatrix} 1 + \sqrt{3} & 1 & 1 \\ 1 & +\sqrt{3} & -2 \\ 1 & -1 & 1 + \sqrt{3} \end{bmatrix}$. The matrix re-

duces to $\begin{bmatrix} 1 + \sqrt{3} & 1 & 1 \\ 0 & 2 + \sqrt{3} & 2\sqrt{3} + 1 \\ 0 & 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_3 = \begin{bmatrix} -1 \\ \sqrt{3} \\ 1 \end{bmatrix}$.

c) $\det(A - \lambda I) = -\lambda^3 + 48\lambda - 128 = -(\lambda + 8)(\lambda - 4)^2$ so the eigenvalues are $\lambda_1 = -8$ and $\lambda_2 = 4$. To find the eigenvector for $\lambda_1 = -8$ we must

solve the homogeneous equation with matrix $\begin{bmatrix} 15 & -9 & -15 \\ 0 & 12 & 0 \\ 3 & -9 & -3 \end{bmatrix}$. The matrix

reduces to $\begin{bmatrix} 15 & -9 & -15 \\ 0 & 12 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. To find

the eigenvector(s) for $\lambda_2 = 4$ we must solve the homogeneous equation with matrix $\begin{bmatrix} 3 & -9 & -15 \\ 0 & 4 & 0 \\ 3 & -9 & -15 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} 3 & -9 & -15 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Thus

there are two eigenvectors corresponding to this eigenvalues, for example $\mathbf{x}_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{x}_3 = \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix}$.

d) $\det(A - \lambda I) = -\lambda^3 + \lambda^2 + \lambda - 1 = -(\lambda + 1)(\lambda - 1)^2$ so the eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 1$. To find the eigenvector for $\lambda_1 = -1$ we must

solve the homogeneous equation with matrix $\begin{bmatrix} 32 & -100 & 70 \\ 18 & -58 & 42 \\ 12 & -40 & 30 \end{bmatrix}$. The matrix

reduces to $\begin{bmatrix} 32 & -100 & 70 \\ 0 & -14 & 21 \\ 0 & 0 & 0 \end{bmatrix}$ and the eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix}$. To find the

eigenvector for $\lambda_2 = 1$ we must solve the homogeneous equation with matrix $\begin{bmatrix} 30 & -100 & 70 \\ 18 & -60 & 42 \\ 12 & -40 & 28 \end{bmatrix}$. The matrix reduces to $\begin{bmatrix} 30 & -100 & 70 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Thus there are

two eigenvectors corresponding to this eigenvalues, for example $\mathbf{x}_2 = \begin{bmatrix} 7 \\ 0 \\ -3 \end{bmatrix}$

and $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 7 \\ 10 \end{bmatrix}$.

Solution 6.6

$$\begin{aligned} \det \begin{bmatrix} p_{11} - \lambda & 1 - p_{22} \\ 1 - p_{11} & p_{22} - \lambda \end{bmatrix} &= (p_{11} - \lambda)(p_{22} - \lambda) - (1 - p_{11})(1 - p_{22}) \\ &= p_{11}p_{22} - (p_{11} + p_{22})\lambda + \lambda^2 - (1 - (p_{11} + p_{22}) + p_{11}p_{22}) \\ &= \lambda^2 - (p_{11} + p_{22})\lambda + (p_{11} + p_{22}) - 1 = 0 \end{aligned}$$

Let $c = p_{11} + p_{22}$, then

$$\begin{aligned}\lambda^2 - c\lambda + c - 1 &= 0 \\ \lambda &= \frac{c \pm \sqrt{c^2 - 4(c-1)}}{2} \\ &= \frac{c \pm \sqrt{c^2 - 4c + 4}}{2} \\ &= \frac{c \pm (c-2)}{2}\end{aligned}$$

Since p_{11} and p_{22} are in the interval $[0, 1]$, we have that $c \in [0, 2]$.

The first eigenvalue is $\lambda_1 = \frac{c+c-2}{2} = c-1$, therefore $\lambda_1 \in [-1, 1]$.

The second eigenvalue is $\lambda_2 = \frac{c-(c-2)}{2} = 1$.

Solution 6.7

$$\begin{aligned}\det(A - \lambda I) &= \det \begin{bmatrix} -\lambda & 1 & -1 \\ 5 & -\lambda & 1 \\ 0 & 1 & -1 - \lambda \end{bmatrix} \\ &= -\lambda \det \begin{bmatrix} -\lambda & 1 \\ 1 & -1 - \lambda \end{bmatrix} - 5 \det \begin{bmatrix} 1 & -1 \\ 1 & -1 - \lambda \end{bmatrix} \\ &= -\lambda(\lambda + \lambda^2 - 1) - 5(-\lambda) = -\lambda^3 - \lambda^2 + \lambda + 5\lambda = 0 \\ &= \lambda(\lambda^2 + \lambda - 6) = \lambda(\lambda - 2)(\lambda + 3)\end{aligned}$$

$$\Rightarrow \lambda = 0, 2, -3.$$

To find the eigenvector corresponding to $\lambda_1 = 0$, we find the RREF of the extended matrix $(A - \lambda_1 I, \vec{0})$, keeping in mind that one of the variables is a free variable.

$$(A - \lambda_1 I, \vec{0}) = \left[\begin{array}{ccc|c} 0 & 1 & -1 & 0 \\ 5 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{5} & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]; \quad \vec{v}_{\lambda_1} = \alpha \begin{bmatrix} -1 \\ 5 \\ 5 \end{bmatrix}.$$

Now, for the eigenvector corresponding to $\lambda_2 = 2$:

$$(A - \lambda_2 I, \vec{0}) = \left[\begin{array}{ccc|c} -2 & 1 & -1 & 0 \\ 5 & -2 & 1 & 0 \\ 0 & 1 & -3 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]; \quad \vec{v}_{\lambda_2} = \alpha \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}.$$

And for the eigenvector corresponding to $\lambda_3 = -3$:

$$(A - \lambda_3 I, \vec{0}) = \left[\begin{array}{ccc|c} 3 & 1 & -1 & 0 \\ 5 & 3 & 1 & 0 \\ 0 & 1 & 2 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]; \quad \vec{v}_{\lambda_3} = \alpha \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

Solution 6.8

$$\begin{aligned}
\det(A - \lambda I) &= \det \begin{bmatrix} 2 - \lambda & 0 & 1 \\ 0 & 2 - \lambda & 1 \\ 1 & 0 & 2 - \lambda \end{bmatrix} \\
&= (2 - \lambda) \det \begin{bmatrix} 2 - \lambda & 1 \\ 0 & 2 - \lambda \end{bmatrix} + \det \begin{bmatrix} 0 & 1 \\ 2 - \lambda & 1 \end{bmatrix} \\
&= (2 - \lambda)^3 - (2 - \lambda) = (2 - \lambda) ((2 - \lambda)^2 - 1) = 0 \\
&= (2 - \lambda)(\lambda^2 - 4\lambda + 3) = (\lambda - 2)(\lambda - 1)(\lambda - 3)
\end{aligned}$$

$$\Rightarrow \lambda = 1, 2, 3.$$

To find the eigenvector corresponding to $\lambda_1 = 1$, we again find the RREF of the extended matrix $(A - \lambda_1 I, \vec{0})$, keeping in mind that one of the variables is a free variable.

$$(A - \lambda_1 I, \vec{0}) = \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]; \quad \vec{v}_{\lambda_1} = \alpha \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

Now, for the eigenvector corresponding to $\lambda_2 = 2$:

$$(A - \lambda_2 I, \vec{0}) = \left[\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]; \quad \vec{v}_{\lambda_2} = \alpha \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

And for the eigenvector corresponding to $\lambda_3 = 3$:

$$(A - \lambda_3 I, \vec{0}) = \left[\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]; \quad \vec{v}_{\lambda_3} = \alpha \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Solution 6.9 The answer is **yes**. The two eigenvectors and eigenvalues define six equations. There are nine unknowns -the entries of M- however, there is an extra condition; the rank of M is two. There are several solutions, here we present one.

Let's try

$$M = \begin{bmatrix} a_{11} & 0 & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix}.$$

We have here six unknowns and six equations, so there is a chance to find a solution. If the two non-zero column vectors are independent, then we are done, it has rank 2.

$$\begin{bmatrix} a_{11} & 0 & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}, \text{ and } \begin{bmatrix} a_{11} & 0 & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -2 \\ -1 \end{bmatrix}$$

$$a_{11} + 3a_{13} = -1, \text{ and } 3a_{11} + a_{13} = -3$$

From this we have $a_{13} = 0$, and $a_{11} = -1$.

$$a_{21} + 3a_{23} = -2, \text{ and } 3a_{21} + a_{23} = -2$$

From this we have that $a_{21} = a_{23} = -\frac{1}{2}$.

$$a_{131} + 3a_{33} = -3, \text{ and } 3a_{31} + a_{33} = -1$$

This is similar to the first pair: $a_{33} = -1$, and $a_{31} = 0$.

The resulting M matrix is

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

M has rank 2, and

$$\begin{bmatrix} -1 & 0 & 0 \\ -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}, \text{ and } \begin{bmatrix} -1 & 0 & 0 \\ -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -2 \\ -1 \end{bmatrix}$$

Solution 6.10

$$\begin{aligned} \det \begin{bmatrix} 2-\lambda & 3 \\ -2 & -1-\lambda \end{bmatrix} &= (2-\lambda)(-1-\lambda) + 6 = -2 - 2\lambda + \lambda + \lambda^2 + 6 \\ &= \lambda^2 - \lambda + 4 = \left(\lambda - \frac{1}{2} + i\frac{\sqrt{15}}{2}\right) \left(\lambda - \frac{1}{2} - i\frac{\sqrt{15}}{2}\right) = 0 \\ \Rightarrow \lambda &= \frac{1}{2} \pm i\frac{\sqrt{15}}{2} \end{aligned}$$

Solution 6.11

a) Find values for a and b such that $A^2 = A$

$$\begin{bmatrix} a & i \\ i & b \end{bmatrix} \begin{bmatrix} a & i \\ i & b \end{bmatrix} = \begin{bmatrix} a & i \\ i & b \end{bmatrix} \quad \Rightarrow \quad \begin{array}{|l|l|} \hline a^2 - 1 = a & ai + bi = i \\ \hline ai + bi = i & -1 + b^2 = b \\ \hline \end{array}$$

Hence, $a^2 - a - 1 = 0$, and $a + b = 1$. Upon solving the system we get

$$a = \frac{1 \pm \sqrt{5}}{2}, \quad b = \frac{1 \mp \sqrt{5}}{2}$$

b) If $A^3 = A$, then we can try

$$A^{-1}A^3 = A^{-1}A \quad \Rightarrow \quad A^2 = I$$

$$\begin{bmatrix} a & i \\ i & b \end{bmatrix} \begin{bmatrix} a & i \\ i & b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \Rightarrow \quad \begin{array}{|l} a^2 - 1 = 1 & ai + bi = 0 \\ ai + bi = 0 & -1 + b^2 = 1 \end{array}$$

Hence, $a^2 = 2$, and $a = -b$. We get

$$a = \pm\sqrt{2}, \quad b = \mp\sqrt{2}$$

$$\begin{bmatrix} \sqrt{2} & i \\ i & -\sqrt{2} \end{bmatrix}^2 = I \quad \Rightarrow \quad A^3 = A; \quad A^4 = I \neq A$$

- c) From the previous solution, we want $A^8 = I \neq A$. If $A^4 = I$, then $A^4 \cdot A^4 = A^8 = I$, hence $A^9 = A$.

Solution 6.12

a)

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

The two eigenvalues are $2 + i$ and $2 - i$. The characteristic polynomial is

$$\begin{aligned} (a - \lambda)(b - \lambda) + 1 &= 0 \\ \lambda^2 - (a + b)\lambda + ab + 1 &= 0 \\ \lambda_{1,2} &= \frac{a + b \pm \sqrt{(a + b)^2 - 4ab - 4}}{2} \\ \lambda_{1,2} &= \frac{a + b}{2} \pm \sqrt{\left(\frac{a + b}{2}\right)^2 - 1} \end{aligned}$$

$\lambda_{1,2}$ should be $2 \pm i$. This is only possible if $\left(\frac{a+b}{2}\right) = 0$, hence $a = b$, and $a + b = 4$. Therefore, $a = b = 2$, and the matrix is

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

- b) Now, to find the eigenvectors, we now that the eigenvalues are $\lambda_{1,2} = 2 + i, 2 - i$. For the first eigenvector we have

$$(A - \lambda_1 I)\vec{v}_1 = \begin{bmatrix} -i & i \\ i & -i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{0}$$

Hence $x_1 = x_2$, and $v_1 = \alpha(1, 1)^T$.

For the second eigenvector we have

$$(A - \lambda_2 I)\vec{v}_2 = \begin{bmatrix} i & i \\ i & i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{0}$$

Hence $x_1 = -x_2$, and $v_2 = \alpha(1, -1)^T$.

Solution 6.13

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

The characteristic polynomial is $\lambda^3 - \frac{1}{2}\lambda^2 - \frac{1}{2}\lambda$. From this, the eigenvalues are found to be $\lambda_{1,2,3} = 0, 1, -\frac{1}{2}$.

To find the eigenvector associated with λ_1 , we find the reduced row-echelon form of the matrix $(P - \lambda_1 I)$:

$$\begin{aligned} \left[\begin{array}{ccc|c} 0 & \frac{1}{4} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & 0 & 0 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 1 & 0 & 0 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & -1 & -2 & 0 \end{array} \right] \\ &\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] &\rightarrow \vec{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \end{aligned}$$

To find the eigenvector associated with λ_2 , we find the RREF of the matrix $(P - \lambda_2 I)$:

$$\begin{aligned} \left[\begin{array}{ccc|c} -1 & \frac{1}{4} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & -1 & 0 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|c} 1 & -\frac{1}{4} & -\frac{1}{2} & 0 \\ 0 & -\frac{3}{4} & \frac{3}{2} & 0 \\ 0 & \frac{3}{4} & -\frac{3}{2} & 0 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & -1 & -2 & 0 \end{array} \right] \\ &\rightarrow \vec{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \end{aligned}$$

Finally, to find the eigenvector associated with λ_3 , we find the RREF of the matrix $(P - \lambda_3 I)$:

$$\begin{aligned} \left[\begin{array}{ccc|c} \frac{1}{2} & \frac{1}{4} & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} & 0 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|c} 1 & \frac{1}{2} & 1 & 0 \\ 0 & \frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] &\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & -2 & 0 \end{array} \right] \\ &\rightarrow \vec{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

We can express $x^{(0)}$ as a linear combination of the eigenvectors:

$$x^{(0)} = \frac{1}{4} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \frac{1}{4} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

From the notes, we know that for $n > 0$ $x^{(n)} = P^n x^{(0)}$, therefore the zero eigenvalue and the $-1/2$ disappear for large n , and we have that

$$\lim_{n \rightarrow \infty} x^{(n)} = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}$$

Solution 6.14

$$P = \begin{bmatrix} \frac{1}{4} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} & b \\ \frac{1}{4} & \frac{1}{3} & a \end{bmatrix}$$

We know that $a \geq 0$, $b \geq 0$, and that $a + b = 1/2$. If $a > 0$, and $b > 0$, by theorem 6.1 we know that the existence of an equilibrium probability is guaranteed. There are two cases left to be analyzed:

i) $a = 0$ and $b = 1/2$

ii) $a = 1/2$ and $b = 0$

In both cases one value of the matrix is zero, so we can't apply theorem 6.1 directly. On the other hand, P^2 for both cases is

i)

$$P^2 = \begin{bmatrix} \frac{17}{48} & \frac{13}{36} & \frac{7}{24} \\ \frac{5}{12} & \frac{9}{36} & \frac{5}{24} \\ \frac{11}{48} & \frac{7}{36} & \frac{17}{24} \end{bmatrix}$$

ii)

$$P^2 = \begin{bmatrix} \frac{17}{48} & \frac{13}{36} & \frac{3}{8} \\ \frac{7}{24} & \frac{5}{18} & \frac{1}{4} \\ \frac{17}{48} & \frac{13}{36} & \frac{3}{8} \end{bmatrix}$$

with all positive entries in both cases, so there is always an equilibrium probability.

Now, for the case $a = b = 1/4$ we have to find the aforementioned equilibrium probability vector. From theorem 6.1, the vector will be the eigenvector corresponding to the eigenvalue $\lambda = 1$, with all its entries summing to one. We find the eigenvalue by finding the reduced row-echelon form of the extended matrix $(P - \lambda I, \vec{0})$:

$$P - I = \begin{bmatrix} -\frac{3}{4} & \frac{1}{3} & \frac{1}{2} & | & 0 \\ \frac{1}{2} & -\frac{2}{3} & \frac{1}{4} & | & 0 \\ \frac{1}{4} & \frac{1}{3} & -\frac{3}{4} & | & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 0 & -5 & | & 0 \\ 0 & 16 & -21 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

$$\rightarrow \vec{v} = \alpha \begin{bmatrix} 20 \\ 21 \\ 16 \end{bmatrix}$$

Since the entries of \vec{v} have to sum to one, we get that the equilibrium probability vector is

$$\frac{1}{57} \begin{bmatrix} 20 \\ 21 \\ 16 \end{bmatrix}$$

Solution 6.15 The eigenvalues and vectors are $\lambda_1 = 2$ $\mathbf{x}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ and $\lambda_2 = 6$ $\mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Thus the general solution is

$$c_1 e^{2t} \begin{bmatrix} -2 \\ 1 \end{bmatrix} + c_2 e^{6t} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

To satisfy the initial condition we need

$$c_1 \begin{bmatrix} -2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

or

$$\begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

which has solution

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

so the solution is

$$-2e^{2t} \begin{bmatrix} -2 \\ 1 \end{bmatrix} + 3e^{6t} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Solution 6.16 The eigenvalues and vectors are $\lambda_1 = 1 + 2i$ $\mathbf{x}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}$ and $\lambda_2 = 1 - 2i$ $\mathbf{x}_2 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$. Thus the general solution in complex form is

$$c_1 e^{(1+2i)t} \begin{bmatrix} i \\ 1 \end{bmatrix} + c_2 e^{(1-2i)t} \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

To find two real solution we take the real and imaginary parts of

$$\begin{aligned} e^{(1+2i)t} \begin{bmatrix} i \\ 1 \end{bmatrix} &= \begin{bmatrix} i e^{(1+2i)t} \\ e^{(1+2i)t} \end{bmatrix} \\ &= e^t \begin{bmatrix} i(\cos(2t) + i \sin(2t)) \\ \cos(2t) + i \sin(2t) \end{bmatrix} \\ &= e^t \begin{bmatrix} -\sin(2t) + i \cos(2t) \\ \cos(2t) + i \sin(2t) \end{bmatrix} \end{aligned}$$

Thus the general solution in real form is

$$a_1 e^t \begin{bmatrix} -\sin(2t) \\ \cos(2t) \end{bmatrix} + a_2 e^t \begin{bmatrix} \cos(2t) \\ \sin(2t) \end{bmatrix}$$

To satisfy the initial condition we need

$$a_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

or

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

which has solution

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

so the solution is

$$e^t \begin{bmatrix} -\sin(2t) + \cos(2t) \\ \cos(2t) + \sin(2t) \end{bmatrix}$$

Solution 6.17 The eigenvalues and vectors are $\lambda_1 = 1$ $\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\lambda_2 = 1 + i$

$\mathbf{x}_2 = \begin{bmatrix} -5 - i \\ -5 - i \\ 2 \end{bmatrix}$ and $\lambda_3 = 1 - i$ $\mathbf{x}_3 = \begin{bmatrix} -5 + i \\ -5 + i \\ 2 \end{bmatrix}$. Thus the general solution in complex form is

$$c_1 e^t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_2 e^{(1+i)t} \begin{bmatrix} -5 - i \\ -5 - i \\ 2 \end{bmatrix} + c_3 e^{(1-i)t} \begin{bmatrix} -5 + i \\ -5 + i \\ 2 \end{bmatrix}$$

The two complex solutions are conjugates of each other. To find real solutions we must find the real and imaginary parts of one of them. These are given by

$e^t \begin{bmatrix} -5 \cos(t) + \sin(t) \\ -5 \cos(t) + \sin(t) \\ 2 \cos(t) \end{bmatrix}$ and $e^t \begin{bmatrix} -5 \sin(t) - \cos(t) \\ -5 \sin(t) - \cos(t) \\ 2 \sin(t) \end{bmatrix}$. Thus the general solution in real form is given by

$$a_1 e^t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + a_2 e^t \begin{bmatrix} -5 \cos(t) + \sin(t) \\ -5 \cos(t) + \sin(t) \\ 2 \cos(t) \end{bmatrix} + a_3 e^t \begin{bmatrix} -5 \sin(t) - \cos(t) \\ -5 \sin(t) - \cos(t) \\ 2 \sin(t) \end{bmatrix}$$

To solve the initial value problem, we must solve

$$a_1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} -5 \\ -5 \\ 2 \end{bmatrix} + a_3 \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

This means solving the system of equations with augmented matrix

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

Reducing this matrix yields

$$\left[\begin{array}{ccc|c} 1 & -5 & -1 & 1 \\ 0 & -5 & -1 & 1 \\ 0 & 0 & -2 & 7 \end{array} \right]$$

Which yields the solution $a_1 = 0$, $a_2 = 1/2$ and $a_3 = -7/2$. So the final answer is

$$\frac{1}{2}e^t \begin{bmatrix} -5\cos(t) + \sin(t) \\ -5\cos(t) + \sin(t) \\ 2\cos(t) \end{bmatrix} - \frac{7}{2}e^t \begin{bmatrix} -5\sin(t) - \cos(t) \\ -5\sin(t) - \cos(t) \\ 2\sin(t) \end{bmatrix} = e^t \begin{bmatrix} \cos(t) + 18\sin(t) \\ \cos(t) + 18\sin(t) \\ \cos(t) - 7\sin(t) \end{bmatrix}$$

Solution 6.18

$$\mathbf{y}'(t) = \begin{bmatrix} 1 & 1 \\ 5 & 1 \end{bmatrix} \mathbf{y}(t)$$

The characteristic polynomial is $(1 - \lambda)^2 - 5 = 0$, hence $\lambda_{1,2} = 1 \pm \sqrt{5}$

We need to find the eigenvectors:

$$\begin{bmatrix} -\sqrt{5} & 1 \\ 5 & -\sqrt{5} \end{bmatrix} \rightarrow \vec{v}_1 = \begin{bmatrix} 1 \\ \sqrt{5} \end{bmatrix}; \quad \begin{bmatrix} \sqrt{5} & 1 \\ 5 & \sqrt{5} \end{bmatrix} \rightarrow \vec{v}_2 = \begin{bmatrix} 1 \\ -\sqrt{5} \end{bmatrix}$$

The general solution is given by

$$\begin{aligned} \mathbf{y}(t) &= c_1 e^{\lambda_1 t} \vec{v}_1 + c_2 e^{\lambda_2 t} \vec{v}_2 \\ &= c_1 e^{(1+\sqrt{5})t} \begin{bmatrix} 1 \\ \sqrt{5} \end{bmatrix} + c_2 e^{(1-\sqrt{5})t} \begin{bmatrix} 1 \\ -\sqrt{5} \end{bmatrix} \end{aligned}$$

Or, equivalently

$$\begin{aligned} y_1(t) &= c_1 e^{(1+\sqrt{5})t} + c_2 e^{(1-\sqrt{5})t} \\ y_2(t) &= c_1 \sqrt{5} e^{(1+\sqrt{5})t} - c_2 \sqrt{5} e^{(1-\sqrt{5})t} \end{aligned}$$

Solution 6.19

$$\mathbf{y}'(t) = \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix} \mathbf{y}(t)$$

The characteristic polynomial is $\lambda^2 - 2\lambda - 2 = 0$, hence $\lambda_{1,2} = 1 \pm \sqrt{3}$

We need to find the eigenvectors:

$$\begin{bmatrix} 2 - \sqrt{3} & -1 \\ 1 & -2 - \sqrt{3} \end{bmatrix} \rightarrow \vec{v}_1 = \begin{bmatrix} 2 + \sqrt{3} \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 + \sqrt{3} & -1 \\ 1 & -2 + \sqrt{3} \end{bmatrix} \rightarrow \vec{v}_2 = \begin{bmatrix} 2 - \sqrt{3} \\ 1 \end{bmatrix}$$

The general solution is given by

$$\mathbf{y}(t) = c_1 e^{(1+\sqrt{3})t} \begin{bmatrix} 2 + \sqrt{3} \\ 1 \end{bmatrix} + c_2 e^{(1-\sqrt{3})t} \begin{bmatrix} 2 - \sqrt{3} \\ 1 \end{bmatrix}$$

The initial condition states that $\mathbf{y}(0) = [0, 1]^T$. Hence

$$\begin{aligned} c_1(2 + \sqrt{3}) + c_2(2 - \sqrt{3}) &= 0 \\ c_1 + c_2 &= 1 \\ c_1 &= \frac{1}{2} - \frac{1}{\sqrt{3}}, \quad c_2 = \frac{1}{2} + \frac{1}{\sqrt{3}} \end{aligned}$$

Or, equivalently

$$\begin{aligned} y_1(t) &= -\frac{1}{2\sqrt{3}}e^{(1+\sqrt{3})t} + \frac{1}{2\sqrt{3}}e^{(1-\sqrt{3})t} \\ y_2(t) &= \left(\frac{1}{2} - \frac{1}{\sqrt{3}}\right)e^{(1+\sqrt{3})t} - \left(\frac{1}{2} + \frac{1}{\sqrt{3}}\right)e^{(1-\sqrt{3})t} \end{aligned}$$

Solution 6.20

$$\mathbf{y}'(t) = \begin{bmatrix} 3 & -1 \\ 7 & 1 \end{bmatrix} \mathbf{y}(t)$$

The characteristic polynomial is $\lambda^2 + 2\lambda + 4 = 0$, hence $\lambda_{1,2} = -1 \pm i\sqrt{3}$

Finding the eigenvectors:

$$\left[\begin{array}{cc|c} -2 - i\sqrt{3} & -1 & 0 \\ 7 & 2 - i\sqrt{3} & 0 \end{array} \right] \rightarrow \vec{v}_1 = \begin{bmatrix} -2 + i\sqrt{3} \\ 7 \end{bmatrix}$$

The function below is a solution. We will separate the real and imaginary parts to get two independent solutions:

$$\begin{aligned} e^{(-1+i\sqrt{3})t} \begin{bmatrix} -2 + i\sqrt{3} \\ 7 \end{bmatrix} &= \frac{1}{e^t} e^{i\sqrt{3}t} \begin{bmatrix} -2 + i\sqrt{3} \\ 7 \end{bmatrix} \\ &= \frac{1}{e^t} \begin{bmatrix} -2 + i\sqrt{3} \\ 7 \end{bmatrix} (\cos(\sqrt{3}t) + i \sin(\sqrt{3}t)) \\ &= \frac{1}{e^t} \begin{bmatrix} -2 \cos(\sqrt{3}t) - \sqrt{3} \sin(\sqrt{3}t) \\ 7 \cos(\sqrt{3}t) \end{bmatrix} + \dots \\ &\quad \frac{i}{e^t} \begin{bmatrix} \sqrt{3} \cos(\sqrt{3}t) - 2 \sin(\sqrt{3}t) \\ 7 \sin(\sqrt{3}t) \end{bmatrix} \end{aligned}$$

The general solution is

$$\mathbf{y}(t) = \frac{c_1}{e^t} \begin{bmatrix} -2 \cos(\sqrt{3}t) - \sqrt{3} \sin(\sqrt{3}t) \\ 7 \cos(\sqrt{3}t) \end{bmatrix} + \frac{c_2}{e^t} \begin{bmatrix} \sqrt{3} \cos(\sqrt{3}t) - 2 \sin(\sqrt{3}t) \\ 7 \sin(\sqrt{3}t) \end{bmatrix}$$

The initial condition states that $\mathbf{y}(0) = [2, -1]^T$. Hence

$$\begin{aligned} -2c_1 + \sqrt{3}c_2 &= 2 \\ 7c_1 &= -1 \\ c_1 &= -\frac{1}{7}, \quad c_2 = \frac{12}{7\sqrt{3}} \end{aligned}$$

Solution 6.21 If A is a 3×3 matrix with real entries, then $\det(A - \lambda I)$ is a polynomial of the form $-\lambda^3 + a\lambda^2 + b\lambda + c$. If you plot the graph of this polynomial, it tends to infinity for large positive λ and to negative infinity for large negative λ . Thus there must be a place where the polynomial crosses the real axis, that is, a real zero. This means that there is always at least one

real eigenvalue. The other two eigenvalues can be either both real, or complex conjugates (as in the previous problem).

Solution 6.22 We have to solve the equation

$$\mathbf{x}'(t) = \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix} \mathbf{x}(t)$$

with

$$\mathbf{x}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The eigenvalues and eigenvectors of the matrix are $\lambda_1 = -1 + i$, $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$, $\lambda_2 = -1 - i$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}$. Real solutions are the real and imaginary parts of

$$e^{(-1+i)t} \begin{bmatrix} 1 \\ -i \end{bmatrix},$$

which are

$$e^{-t} \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$

and

$$e^{-t} \begin{bmatrix} \sin(t) \\ -\cos(t) \end{bmatrix}$$

Thus the general solution is

$$\mathbf{x}(t) = a_1 e^{-t} \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} + a_2 e^{-t} \begin{bmatrix} \sin(t) \\ -\cos(t) \end{bmatrix}.$$

To satisfy the initial conditions, we need

$$a_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

so $a_1 = 1$, $a_2 = -1$ and the solution is

$$e^{-t} \begin{bmatrix} \cos(t) - \sin(t) \\ \sin(t) + \cos(t) \end{bmatrix}$$

Solution 6.23 The current is the same through each component. The voltage across the resistor is IR , so if V_L is the voltage across the inductor, then $V_L + V + IR = 0$ so $V_L = -V - IR$. The equations are therefore

$$\begin{aligned} I' &= \frac{1}{L} V_L = -\frac{R}{L} I - \frac{1}{L} V \\ V' &= \frac{1}{C} I \end{aligned}$$

or

$$\begin{bmatrix} I \\ V \end{bmatrix}' = \begin{bmatrix} -\frac{R}{L} & -\frac{1}{L} \\ \frac{1}{C} & 0 \end{bmatrix} \begin{bmatrix} I \\ V \end{bmatrix}$$

The characteristic polynomial (i.e., $\det(A - \lambda I)$) is $\lambda^2 + (R/L)\lambda + 1/(LC)$, so the eigenvalues are $\lambda = (-R/L \pm \sqrt{R^2/L^2 - 4/(LC)})/2$. Oscillations occur when the eigenvalues are complex, so, if $R^2/L^2 < 4/(LC)$ or $R < 2\sqrt{L/C}$. If we set all the values to 1 the equation becomes

$$\begin{bmatrix} I \\ V \end{bmatrix}' = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} I \\ V \end{bmatrix}$$

and the initial conditions are

$$\begin{bmatrix} I(0) \\ V(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The eigenvalues and vectors are $\lambda_1 = (-1 + i\sqrt{3})/2$, $\mathbf{x}_1 = \begin{bmatrix} (-1 + i\sqrt{3})/2 \\ 1 \end{bmatrix}$, $\lambda_2 = (-1 - i\sqrt{3})/2$, $\mathbf{x}_2 = \begin{bmatrix} (-1 - i\sqrt{3})/2 \\ 1 \end{bmatrix}$. Real solutions are given by the real and imaginary parts of

$$\begin{aligned} e^{(-1+i\sqrt{3})t/2} \begin{bmatrix} (-1+i\sqrt{3})/2 \\ 1 \end{bmatrix} &= e^{-t/2} \begin{bmatrix} (-1/2 + i\sqrt{3}/2)(\cos(\sqrt{3}t/2) + i\sin(\sqrt{3}t/2)) \\ \cos(\sqrt{3}t/2) + i\sin(\sqrt{3}t/2) \end{bmatrix} \\ &= e^{-t/2} \begin{bmatrix} -\cos(\sqrt{3}t/2)/2 - \sqrt{3}\sin(\sqrt{3}t/2)/2 + i(-\sin(\sqrt{3}t/2)/2 + \sqrt{3}\cos(\sqrt{3}t/2)/2) \\ \cos(\sqrt{3}t/2) + i\sin(\sqrt{3}t/2) \end{bmatrix} \end{aligned}$$

So the general solution is

$$a_1 e^{-t/2} \begin{bmatrix} -\cos(\sqrt{3}t/2)/2 - \sqrt{3}\sin(\sqrt{3}t/2)/2 \\ \cos(\sqrt{3}t/2) \end{bmatrix} + a_2 e^{-t/2} \begin{bmatrix} -\sin(\sqrt{3}t/2)/2 + \sqrt{3}\cos(\sqrt{3}t/2)/2 \\ \sin(\sqrt{3}t/2) \end{bmatrix}$$

To satisfy the initial conditions, we need

$$a_1 \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix} + a_2 \begin{bmatrix} \sqrt{3}/2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

This has solution $a_1 = \sqrt{3}/3$, $a_2 = -2 - \sqrt{3}/3$.

Solution 6.24 The equivalent 2×2 system is

$$\mathbf{z}' = \begin{bmatrix} 0 & 1 \\ -6 & 5 \end{bmatrix} \mathbf{z}$$

with

$$\mathbf{z}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The eigenvalues and eigenvectors are $\lambda_1 = 2$, $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\lambda_2 = 3$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

The general solution is

$$c_1 e^{2t} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 e^{3t} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

The initial conditions are satisfied if

$$c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

which has solution $c_1 = 3$, $c_2 = -2$. So $y(t) = z_1(t) = 3e^{2t} - 2e^{3t}$. To solve this directly, plug in $e^{\lambda t}$. This solves the equation if $\lambda^2 - 5\lambda + 6 = 0$, or $\lambda = 2$ or 3 . So the general solution is $y(t) = c_1 e^{2t} + c_2 e^{3t}$. Satisfying the initial conditions $y(0) = 1$ and $y'(0) = 0$ leads to $c_1 = 3$ and $c_2 = -2$ as above.

Solution 6.25 The equivalent 2×2 system is

$$\mathbf{z}' = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \mathbf{z}$$

with

$$\mathbf{z}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The eigenvalues and eigenvectors are $\lambda_1 = -1/2 + i\sqrt{3}/2$, $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1/2 + i\sqrt{3}/2 \end{bmatrix}$, $\lambda_2 = -1/2 - i\sqrt{3}/2$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1/2 - i\sqrt{3}/2 \end{bmatrix}$. Real solutions are the real and imaginary parts of

$$e^{(-1/2 + i\sqrt{3}/2)t} \begin{bmatrix} 1 \\ -1/2 + i\sqrt{3}/2 \end{bmatrix}$$

which leads to the general solution

$$a_1 e^{-t/2} \begin{bmatrix} \cos \\ -(1/2) \cos - (\sqrt{3}/2) \sin \end{bmatrix} + a_2 e^{-t/2} \begin{bmatrix} \sin \\ (\sqrt{3}/2) \cos - (1/2) \sin \end{bmatrix}$$

(the sines and cosines are evaluated at $\sqrt{3}t/2$). Initial conditions require

$$a_1 \begin{bmatrix} 1 \\ -(1/2) \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ (\sqrt{3}/2) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

which has solution $a_1 = 1$ and $a_2 = \sqrt{3}/3$. Thus the solution is $y(t) = z_1(t) = e^{-t/2}(\cos(\sqrt{3}t/2) + (\sqrt{3}/3)\sin(\sqrt{3}t/2))$. If we try to solve directly by substituting $e^{\lambda t}$, we get $\lambda = -1/2 \pm i\sqrt{3}/2$. The real and imaginary parts of $e^{\lambda t}$ are $e^{-t/2}(\cos(\sqrt{3}t/2))$ and $e^{-t/2}(\sin(\sqrt{3}t/2))$, so we obtain the general solution $y(t) = a_1 e^{-t/2}(\cos(\sqrt{3}t/2) + a_2 e^{-t/2} \sin(\sqrt{3}t/2))$. Choosing a_1 and a_2 to solve the initial condition yields the same answer as above.

Solution 6.26 Set $z_1(t) = y(t)$, $z_2(t) = y'(t)$ and $z_3(t) = y''(t)$. Then $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$

solves the equation

$$\mathbf{z}' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -c & -b & -a \end{bmatrix} \mathbf{z}$$

Solution 6.27 The general solution is

$$(a_1 \cos(t) + b_1 \sin(t)) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (a_2 \cos(2t) + b_2 \sin(2t)) \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + (a_3 \cos(3t) + b_3 \sin(3t)) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

To satisfy the initial conditions, we need

$$a_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + a_3 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

which has solution $a_1 = 1$, $a_2 = 0$ and $a_3 = 2$, and

$$b_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 2b_2 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 3b_3 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

which has solution $b_1 = 1/2$, $2b_2 = -1/2$ and $3b_3 = 1$, so that $b_1 = 1/2$, $b_2 = -1/4$ and $b_3 = 1/3$

Solution 6.28 Going through the analysis with two springs and weights, we end up with

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

and

$$\mathbf{b} = \begin{bmatrix} 10 \\ 11 \end{bmatrix}$$

Then

$$\mathbf{x}_{\text{eq}} = K^{-1}\mathbf{b} = \begin{bmatrix} 21 \\ 31 \end{bmatrix}$$

so if we define $\mathbf{y} = \mathbf{x} - \mathbf{x}_{\text{eq}}$, we obtain initial conditions

$$\mathbf{y}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$\mathbf{y}'(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The eigenvalues of K are $\lambda_1 = (-3 + \sqrt{5})/2$ and $\lambda_2 = (-3 - \sqrt{5})/2$ with eigenvectors $\mathbf{x}_1 = \begin{bmatrix} (-1 + \sqrt{5})/2 \\ 1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} (-1 - \sqrt{5})/2 \\ 1 \end{bmatrix}$. If we set $\mu_1 = \sqrt{-\lambda_1}$ and $\mu_2 = \sqrt{-\lambda_2}$, then the general solution is

$$\mathbf{y}(t) = (a_1 \cos(\mu_1 t) + b_1 \sin(\mu_1 t))\mathbf{x}_1 + (a_2 \cos(\mu_2 t) + b_2 \sin(\mu_2 t))\mathbf{x}_2$$

The initial condition are satisfied if

$$a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$\mu_1 b_1 \mathbf{x}_1 + \mu_2 b_2 \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

These can be solved by inverting 2×2 matrices. The answers are

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} -3/2 + \sqrt{5}/2 \\ 3/2 + \sqrt{5}/2 \end{bmatrix}$$

and

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} (-3/2 + \sqrt{5}/2)/\mu_1 \\ (3/2 + \sqrt{5}/2)/\mu_2 \end{bmatrix}$$