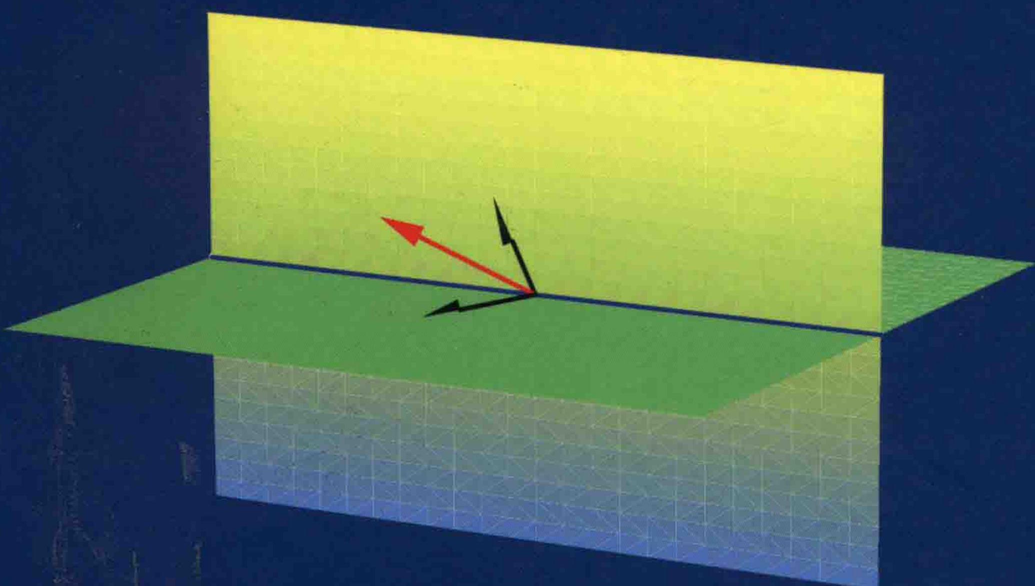
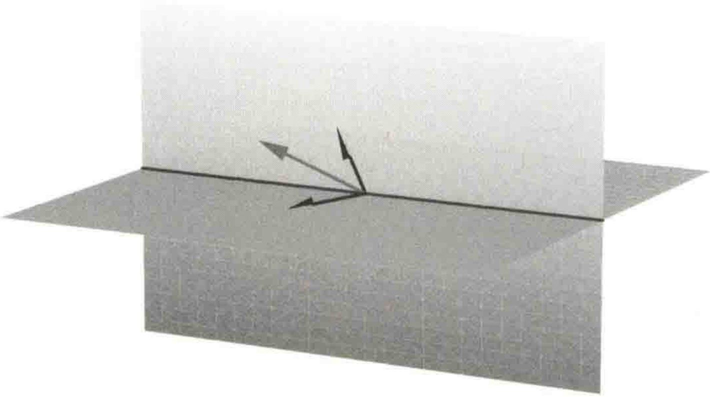
LINEAR ALGEBRA AS AN INTRODUCTION TO ABSTRACT MATHEMATICS

Isaiah Lankham
Bruno Nachtergaele
Anne Schilling



World Scientific



LINEAR ALGEBRA AS AN INTRODUCTION TO ABSTRACT MATHEMATICS

Isaiah Lankham
California State University, East Bay, USA

Bruno Nachtergaele
University of California, Davis, USA

Anne Schilling
University of California, Davis, USA

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Published by

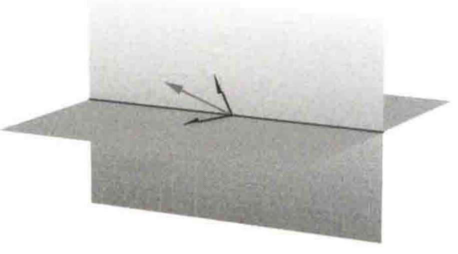
World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

LINEAR ALGEBRA AS AN INTRODUCTION TO ABSTRACT MATHEMATICS

ISBN 978-981-4730-35-8
ISBN 978-981-4723-77-0 (pbk)

Printed in Singapore



LINEAR ALGEBRA AS AN INTRODUCTION TO ABSTRACT MATHEMATICS

Preface

"Linear Algebra - As an Introduction to Abstract Mathematics" is an introductory textbook designed for undergraduate mathematics majors and other students who do not shy away from an appropriate level of abstraction. In fact, we aim to introduce abstract mathematics and proofs in the setting of linear algebra to students for whom this may be the first step toward advanced mathematics. Typically, such a student will have taken calculus, though the only prerequisite is suitable mathematical maturity. The purpose of this book is to bridge the gap between more conceptual and computational oriented lower division undergraduate classes and more abstract oriented upper division classes.

The book begins with systems of linear equations and complex numbers, then relates these to the abstract notion of linear maps on finite-dimensional vector spaces, and covers diagonalization, eigenspaces, determinants, and the spectral theorem. Each chapter concludes with both proof-writing and computational exercises.

We wish to thank our many undergraduate students who took MAT67 at UC Davis in the past several years and our colleagues who taught from our lecture notes that eventually became this book. Their comments on earlier drafts were invaluable. This book is dedicated to them and all future students and teachers who use it.

I. Lankham

B. Nachtergaele

A. Schilling

California, October 2015

Contents

Chapter 1

What is Linear Algebra?

1.1 Introduction

This book aims to bridge the gap between the mainly computation-oriented lower division undergraduate classes and the abstract mathematics encountered in more advanced mathematics courses. The goal of this book is threefold:

(1) You will learn **Linear Algebra**, which is one of the most widely used mathematical theories around. Linear Algebra finds applications in virtually every area of mathematics, including multivariate calculus, differential equations, and probability theory. It is also widely applied in fields like physics, chemistry, economics, psychology, and engineering. You are even relying on methods from Linear Algebra every time you use an internet search like Google, the Global Positioning System (GPS), or a cellphone.
(2) You will acquire **computational skills** to solve linear systems of equations, perform operations on matrices, calculate eigenvalues, and find determinants of matrices.
(3) In the setting of Linear Algebra, you will be introduced to **abstraction**. As the theory of Linear Algebra is developed, you will learn how to make and use definitions and how to write proofs.

The exercises for each Chapter are divided into more computation-oriented exercises and exercises that focus on proof-writing.

1.2 What is Linear Algebra?

Linear Algebra is the branch of mathematics aimed at solving systems of linear equations with a finite number of unknowns. In particular, one would like to obtain answers to the following questions:

- **Characterization of solutions:** Are there solutions to a given system of linear equations? How many solutions are there?
- **Finding solutions:** How does the solution set look? What are the solutions?

Linear Algebra is a systematic theory regarding the solutions of systems of linear equations.

Example 1.2.1. Let us take the following system of two linear equations in the two unknowns x_1 and x_2:

$$\left.\begin{aligned} 2x_1 + x_2 &= 0 \\ x_1 - x_2 &= 1 \end{aligned}\right\}.$$

This system has a **unique solution** for $x_1, x_2 \in \mathbb{R}$, namely $x_1 = \frac{1}{3}$ and $x_2 = -\frac{2}{3}$.

The solution can be found in several different ways. One approach is to first solve for one of the unknowns in one of the equations and then to substitute the result into the other equation. Here, for example, we might solve to obtain

$$x_1 = 1 + x_2$$

from the second equation. Then, substituting this in place of x_1 in the first equation, we have

$$2(1 + x_2) + x_2 = 0.$$

From this, $x_2 = -2/3$. Then, by further substitution,

$$x_1 = 1 + \left(-\frac{2}{3}\right) = \frac{1}{3}.$$

Alternatively, we can take a more systematic approach in eliminating variables. Here, for example, we can subtract 2 times the second equation from the first equation in order to obtain $3x_2 = -2$. It is then immediate that $x_2 = -\frac{2}{3}$ and, by substituting this value for x_2 in the first equation, that $x_1 = \frac{1}{3}$.

Example 1.2.2. Take the following system of two linear equations in the two unknowns x_1 and x_2:

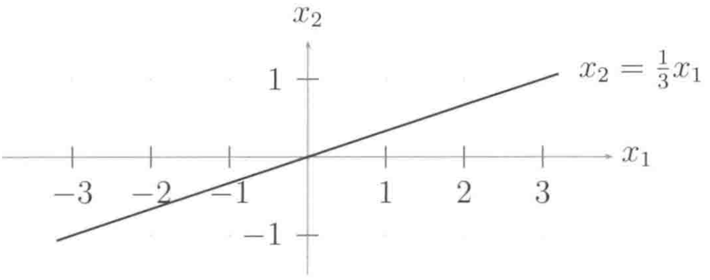
$$\left.\begin{aligned} x_1 + x_2 &= 1 \\ 2x_1 + 2x_2 &= 1 \end{aligned}\right\}.$$

We can eliminate variables by adding -2 times the first equation to the second equation, which results in $0 = -1$. This is obviously a contradiction, and hence this system of equations has **no solution**.

Example 1.2.3. Let us take the following system of one linear equation in the two unknowns x_1 and x_2:

$$x_1 - 3x_2 = 0.$$

In this case, there are **infinitely many** solutions given by the set $\{x_2 = \frac{1}{3}x_1 \mid x_1 \in \mathbb{R}\}$. You can think of this solution set as a line in the Euclidean plane $\mathbb{R}^2$:



In general, a **system of m linear equations in n unknowns** $x_1, x_2, \ldots, x_n$ is a collection of equations of the form

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}\right\}, \tag{1.1}$$

where the a_{ij}'s are the coefficients (usually real or complex numbers) in front of the unknowns x_j, and the b_i's are also fixed real or complex numbers. A **solution** is a set of numbers $s_1, s_2, \ldots, s_n$ such that, substituting $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ for the unknowns, all of the equations in System (1.1) hold. Linear Algebra is a theory that concerns the solutions and the structure of solutions for linear equations. As we progress, you will see that there is a lot of subtlety in fully understanding the solutions for such equations.

1.3 Systems of linear equations

1.3.1 *Linear equations*

Before going on, let us reformulate the notion of a system of linear equations into the language of functions. This will also help us understand the adjective "linear" a bit better. A **function** f is a map

$$f : X \to Y \tag{1.2}$$

from a set X to a set Y. The set X is called the **domain** of the function, and the set Y is called the **target space** or **codomain** of the function. An **equation** is

$$f(x) = y, \tag{1.3}$$

where $x \in X$ and $y \in Y$. (If you are not familiar with the abstract notions of sets and functions, please consult Appendix B.)

Example 1.3.1. Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = x^3 - x$. Then $f(x) = x^3 - x = 1$ is an equation. The domain and target space are both the set of real numbers $\mathbb{R}$ in this case.

In this setting, a system of equations is just another kind of equation.

Example 1.3.2. Let $X = Y = \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ be the Cartesian product of the set of real numbers. Then define the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ as

$$f(x_1, x_2) = (2x_1 + x_2, x_1 - x_2), \tag{1.4}$$

and set $y = (0, 1)$. Then the equation $f(x) = y$, where $x = (x_1, x_2) \in \mathbb{R}^2$, describes the system of linear equations of Example 1.2.1.

The next question we need to answer is, "What is a linear equation?". Building on the definition of an equation, a **linear equation** is any equation defined by a "linear" function f that is defined on a "linear" space (a.k.a. a **vector space** as defined in Section 4.1). We will elaborate on all of this in later chapters, but let us demonstrate the main features of a "linear" space in terms of the example $\mathbb{R}^2$. Take $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$. There are two "linear" operations defined on $\mathbb{R}^2$, namely addition and scalar multiplication:

$$x + y := (x_1 + y_1, x_2 + y_2) \qquad \text{(vector addition)} \tag{1.5}$$

$$cx := (cx_1, cx_2) \qquad \text{(scalar multiplication).} \tag{1.6}$$

A "linear" function on $\mathbb{R}^2$ is then a function f that interacts with these operations in the following way:

$$f(cx) = cf(x) \tag{1.7}$$

$$f(x + y) = f(x) + f(y). \tag{1.8}$$

You should check for yourself that the function f in Example 1.3.2 has these two properties.

1.3.2 *Non-linear equations*

(Systems of) Linear equations are a very important class of (systems of) equations. We will develop techniques in this book that can be used to solve any systems of linear equations. Non-linear equations, on the other hand, are significantly harder to solve. An example is a **quadratic equation** such as

$$x^2 + x - 2 = 0, \tag{1.9}$$

which, for no completely obvious reason, has exactly two solutions $x = -2$ and $x = 1$. Contrast this with the equation

$$x^2 + x + 2 = 0, \tag{1.10}$$

which has no solutions within the set $\mathbb{R}$ of real numbers. Instead, it has two complex solutions $\frac{1}{2}(-1 \pm i\sqrt{7}) \in \mathbb{C}$, where $i = \sqrt{-1}$. (Complex numbers are discussed in more detail in Chapter 2.) In general, recall that the quadratic equation $x^2 + bx + c = 0$ has the two solutions

$$x = -\frac{b}{2} \pm \sqrt{\frac{b^2}{4} - c}.$$

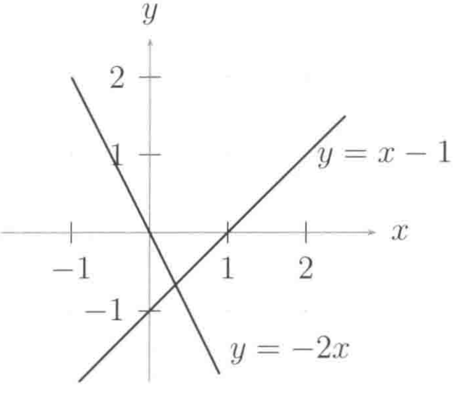
1.3.3 *Linear transformations*

The set $\mathbb{R}^2$ can be viewed as the Euclidean plane. In this context, linear functions of the form $f : \mathbb{R}^2 \to \mathbb{R}$ or $f : \mathbb{R}^2 \to \mathbb{R}^2$ can be interpreted geometrically as "motions" in the plane and are called **linear transformations**.

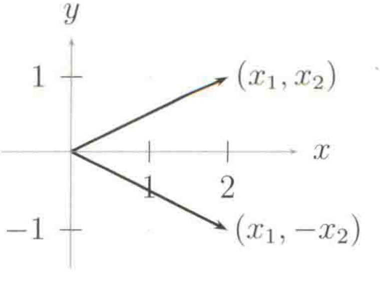
Example 1.3.3. Recall the following linear system from Example 1.2.1:

$$\left.\begin{aligned} 2x_1 + x_2 &= 0 \\ x_1 - x_2 &= 1 \end{aligned}\right\}.$$

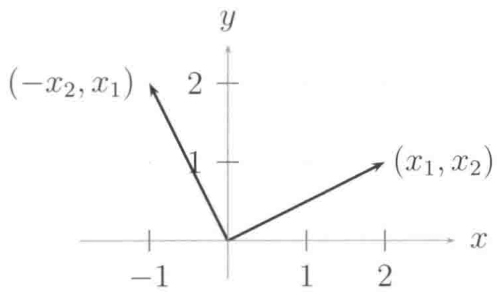
Each equation can be interpreted as a straight line in the plane, with solutions (x_1, x_2) to the linear system given by the set of all points that simultaneously lie on both lines. In this case, the two lines meet in only one location, which corresponds to the unique solution to the linear system as illustrated in the following figure:



Example 1.3.4. The linear map $f(x_1, x_2) = (x_1, -x_2)$ describes the "motion" of reflecting a vector across the x-axis, as illustrated in the following figure:



Example 1.3.5. The linear map $f(x_1, x_2) = (-x_2, x_1)$ describes the "motion" of rotating a vector by 90° counterclockwise, as illustrated in the following figure:



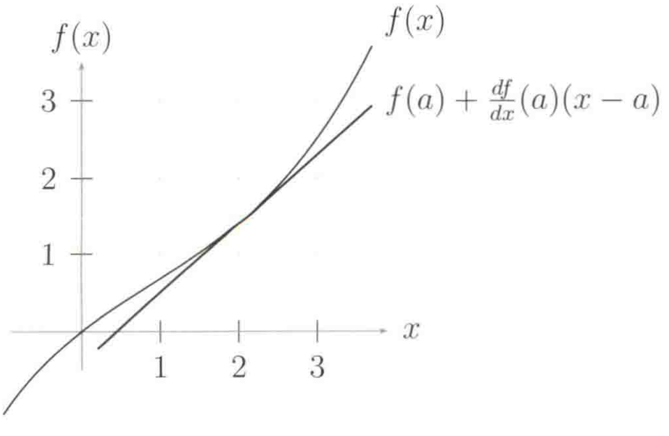
This example can easily be generalized to rotation by any arbitrary angle using Lemma 2.3.2. In particular, when points in $\mathbb{R}^2$ are viewed as complex numbers, then we can employ the so-called polar form for complex numbers in order to model the "motion" of rotation. (Cf. Proof-Writing Exercise 5 on page 20.)

1.3.4 *Applications of linear equations*

Linear equations pop up in many different contexts. For example, you can view the derivative $\frac{df}{dx}(x)$ of a differentiable function $f : \mathbb{R} \to \mathbb{R}$ as a linear approximation of f. This becomes apparent when you look at the Taylor series of the function $f(x)$ centered around the point $x = a$ (as seen in calculus):

$$f(x) = f(a) + \frac{df}{dx}(a)(x - a) + \cdots . \tag{1.11}$$

In particular, we can graph the linear part of the Taylor series versus the original function, as in the following figure:



Since $f(a)$ and $\frac{df}{dx}(a)$ are merely real numbers, $f(a)+\frac{df}{dx}(a)(x-a)$ is a linear function in the single variable x.

Similarly, if $f : \mathbb{R}^n \to \mathbb{R}^m$ is a multivariate function, then one can still view the derivative of f as a form of a linear approximation for f (as seen in a multivariate calculus course).