LINEAR ALGEBRA WITH APPLICATIONS

A THESIS SUBMITTED TO THE
GRADUATE SCHOOL OF APPLIED SCIENCES
OF
NEAR EAST UNIVERSITY

by
BASSMA ABDLRAZG

In Partial Fulfillment of the Requirements for
the Degree of Master of Science
in
Mathematics

NICOSIA, 2016

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, last name:

Signature:

ACKNOWLEDGMENTS

I truly feel very thankful to my supervisor Assoc. Prof. Dr. Burak Sekeroglu for his assistance, guidance and supervision of my thesis. I appreciate his continuous follow up, support and motivation. He was always sharing his time and effort whenever I need him.

I acknowledge Assoc. Prof. Dr. Evren Hincal for his understanding, supporting and being always there for any advice.

I also appreciate NEU Grand Library administration members for offering perfect environment for study, research and their efforts to provide the updated research materials and resources.

I also send my special thanks to my mother for her care, prayers and her passion. I also appreciate my father's continuous support, advice and encouragement. I would also like to say thanks to my Husband “Tarek” for his attention, support and availability when I need him.

Finally, I also have to thank God for everything and for supplying me with patience and supporting me with faith.

ABSTRACT

Linear algebra is a main important part of the mathematics. It is a principal branch of mathematics that is related to mathematical structures closed under the operations of addition and scalar multiplication and that includes the theory of systems of linear equations, matrices, determinants, vector spaces, and linear transformations. Linear algebra, is a mathematical discipline that deals with vectors and matrices and, more generally, with vector spaces and linear transformations. Unlike other parts of mathematics that are frequently invigorated by new ideas and unsolved problems, linear algebra is very well understood. Its value lies in its many applications, from mathematical physics to modern algebra and its usage in the engineering and medical fields such as image processing and analysis.

This thesis is a detailed review and explanation of the linear algebra domain in which all mathematical concepts and structures concerned with linear algebra are discussed. The thesis's main aim is to point out the significant applications of the linear algebra in the medical engineering field. Hence, the eigenvectors and eigenvalues which represent the core of linear algebra are discussed in details in order to show how they can be used in many engineering applications. The principal components analysis is one of the most important compression and feature extraction algorithms used in the engineering field. It is mainly dependent on the calculation and extraction of eigenvalues and eigenvectors that then be used to represent an input; whether it is image or a simple matrix. In this thesis, the use of principal components analysis for the compression of medical images is discussed as an important and novel application of linear algebra.

Keywords: Linear algebra; addition; scalar; multiplication; linear equations; matrices; determinants; vector spaces; linear transformations; image processing; eigenvectors; eigenvalues; principal components analysis; compression

ÖZET

Lineer Cebir matematiğin en önemli parçalarından biridir ve Matematiğin, toplama ve sayıl çarpma gibi işlemlere göre daha kapalı olan Matematiksel yapılar ile ilgili olan ve doğrusal denklem sistemi, matrisler, determinant, vektör uzayları ve lineer dönüşümler teorilerini içeren bir ana bilim dalıdır. Lineer Cebir, vektörler ve matrisleri, daha genel anlamda ise vektör uzayları ve lineer dönüşümleri ele alan bir matematik bilim dalıdır. Matematiğin sıklıkla yeni fikirler ve çözümlenmemiş problemlerle gündemde kalan diğer dallarının aksine, lineer cebir daha anlaşılır bir konumdadır. Lineer Cebirin değeri matematiksel fizikten modern Cebire kadar uzanan birçok uygulama yanında görüntü işleme ve analiz gibi mühendislik ve tıp alanlarında da kullanılmasından kaynaklanmaktadır.

Bu tez, Lineer Cebirle ilgili olan tüm Matematiksel kavramların ve yapıların ele alındığı, ve bu alanla ilgili detaylı bir inceleme ve açıklamadır. Tezin esas amacı, Lineer Cebirin Medikal Mühendislik alanında kullanılan önemli uygulamalarına dikkat çekmektir. Bu nedenle, lineer Cebirin özünü oluşturan özvektörler ve özdeğerlerin birçok mühendislik uygulamasında nasıl kullanılabileceğini göstermek amacıyla detaylıca ele alınmıştır. Ana bileşenler analizi, mühendislik alanında kullanılan en önemli sıkıştırma ve öznitelik çıkarımı algoritmalarından biridir. Bu esasen, daha sonradan bir veriyi temsil edecek olan özdeğerler ve özvektörler çıkarımı ve hesaplanmasına bağlıdır; bir görüntü veya basit bir matris de olabilir. Bu tezde, Lineer Cebirin önemli ve yeni bir uygulaması olarak, ana bileşenler analizinin medikal görüntülerin kompresyonu için kullanılması ele alınmıştır.

Anahtar kelimeler: Lineer cebir; ekleme; sayıl; çarpma; lineer denklemler; matrisler; determinantlar; vektör uzayları; doğrusal dönüşümler; görüntü işleme; özvektörler; özdeğerler; temel bileşenler analizi; sıkıştırma

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Linear algebra is an important course for a diverse number of students for at least two reasons. First, few subjects can claim to have such widespread applications in other areas of mathematics-multi variable calculus, differential equations, and probability, for example-as well as in physics, biology, chemistry, economics, finance, psychology, sociology, and all fields of engineering. Second, this subject presents the student at the sophomore level with an excellent opportunity to learn how to handle abstract concepts.

Linear algebra is one of the most known mathematical disciplines because of its rich theoretical foundations and its many useful applications to science and engineering. Solving systems of linear equations and computing determinants are two examples of fundamental problems in linear algebra that have been studied for a long time ago. Leibnitz found the formula for determinants in 1693, and in 1750 Cramer presented a method for solving systems of linear equations, which is today known as Cramer's Rule. This is the first foundation stone on the development of linear algebra and matrix theory. At the beginning of the evolution of digital computers, the matrix calculus has received very much attention. John von Neumann and Alan Turing were the world-famous pioneers of computer science. They introduced significant contributions to the development of computer linear algebra. In 1947, von Neumann and Goldstine investigated the effect of rounding errors on the solution of linear equations. One year later, Turing [Tur48] initiated a method for factoring a matrix to a product of a lower triangular matrix with an echelon matrix (the factorization is known as LU decomposition). At present, computer linear algebra is broadly of interest. This is due to the fact that the field is now recognized as an absolutely essential tool in many branches of computer applications that require computations which are lengthy and difficult to get right when done by hand, for example: in computer graphics, in geometric modeling, in robotics, etc.

1.2 Aims of Thesis

The motivation for this thesis comes mainly from the purpose to understand the complexity of mathematical problems in linear algebra. Many tasks of linear algebra are recognized usually as elementary problems, but the precise complexity of them was not known for a long time ago.

The aims are of this thesis is to understand the eigenvalues and eigenvectors and to go through some of their applications in the mathematical and engineering areas in order to show their importance and impact.

1.3 Thesis Overview

This thesis is structured as follows:

Chapter 1 is an introduction of the thesis; it presents the aims of thesis as well as the thesis overview.

Chapter 2 introduces the basics of linear algebra. It first introduces the linear algebra as a concept. Then, it discusses the scalars properties such as distributivity, and commutativity etc.. the vectors space mathematical operation are also discussed such as addition, multiplication, and subtraction.

Chapter 3 deals with matrices and their properties. In this chapter we also provide a clear introduction to matrix transformations and an application of the dot product to statistics. This chapter introduces the basic properties of determinants and some of their applications as well as the systems of linear equations

Chapter 4 presents a simple explanation of the linear combinations as well as linear independence.

Chapter 5 presents different types of linear transformation of matrices and also different properties of them.

Chapter 6 considers eigenvalues and eigenvectors. In this chapter we completely solve the diagonalization problem for symmetric matrices in addition to other application of the eigenvalues and eigenvectors such as PCA. In here, a detailed explanation of the PCA is presented. A medical engineering application of the PCA is presented in this chapter in order to point out the importance of the eigenvalues and eigenvectors in engineering applications. **Chapter 7** is a conclusion of the presented thesis.

CHAPTER TWO
LINEAR ALGEBRA BASICS

This chapter reviews the basic concepts and thoughts of linear algebra. It discusses and reviews the scalars and their properties through equations. Moreover, it presents the vectors and their transformations such as multiplication, subtraction etc..

2.1 Introduction to Linear Algebra

Linear Algebra is a standout amongst the most critical fundamental ranges in Mathematics, having at any rate as awesome an effect as Calculus, and to be sure it gives a noteworthy piece of the hardware required to sum up Calculus to vector-esteemed elements of numerous variables. Dissimilar to numerous logarithmic frameworks considered in Mathematics or connected inside or out with it, a hefty portion of the issues concentrated on in Linear Algebra are manageable to precise and even algorithmic arrangements, and this makes them implementable on PCs – this clarifies why so much calculational utilization of PCs includes this sort of polynomial math and why it is so generally utilized. Numerous geometric subjects are examined making utilization of ideas from Linear Algebra, and the thought of a direct change is an arithmetical adaptation of geometric change. At long last, a lot of present day unique variable based math constructs on Linear Algebra and regularly gives solid illustrations of general though (Poole, 2010).

The subject of linear algebra based math can be somewhat clarified by the means of the two terms involving the title. "Linear" is a term you will acknowledge better toward the end of this course, and in reality, achieving this gratefulness could be taken as one of the essential objectives of this course. However until further notice, you can comprehend it to mean anything that is "straight" or "level." For instance in the *xy*-plane you may be acclimated to portraying straight lines (is there some other kind?) as the arrangement of answers for a mathematical statement of the structure $y=mx+b$, where the slant m and the y-capture b are constants that together depict the line. In the event that you have contemplated multivariate analytics, then you will have experienced planes. Living in three measurements, with directions portrayed by triples (x,y,z), they can be depicted as the arrangement of answers for mathematical statements of the structure $ax+by+cz=d$, where a,b,c,d are constants that together focus the plane. While we may depict planes as level, lines in three measurements may be portrayed as linear. From a multivariate analytics course you will review that lines are sets of focuses portrayed by comparisons, for example, $x=3t-4$, $y=-7t+2$, $z=9t$, where t is a parameter that can tackle any worth.

Another perspective of this idea of levelness is to perceive that the arrangements of focuses simply depicted are answers for mathematical statements of a moderately basic structure. These mathematical statements include expansion and duplication just. We will have a requirement for subtraction, and every so often we will isolate, yet for the most part you can depict linear mathematical statements as including just addition and multiplication (Kolman, 1996).

2.2 Scalars

Before examining vectors, first we clarify what is implied by scalars. These are "numbers" of different sorts together with logarithmic operations for consolidating them. The principle cases we will consider are the objective numbers Q, the genuine numbers R and the mind boggling numbers C. Be that as it may mathematicians routinely work with different fields, for example, the limited fields (otherwise called Galois fields) which are essential in coding hypothesis, cryptography and other advanced applications (Rajendra, 1996).

A field of scalars (or only a field) comprises of a set F whose components are called scalars, together with two arithmetical operations, expansion + and augmentation ×, for joining each pair of scalars x, y ∈ F to give new scalars x + y ∈ F and x × y ∈ F. These operations are required to fulfill the accompanying properties which are here and there known as the field

Associativity: For x, y, z ∈ F,

$$(x + y) + z = x + (y + z), \tag{2.1}$$

$$(x \times y) \times z = x \times (y \times z) \tag{2.2}$$

Zero and unity: There are unique and distinct elements 0, 1 ∈ F such that for x ∈ F,

$$x + 0 = x = 0 + x, \tag{2.3}$$

$$x \times 1 = x = 1 \times x. \tag{2.4}$$

Distributivity: For x, y, z ∈ F,

$$(x + y) \times z = x \times z + y \times z, \quad (2.5)$$

$$z \times (x + y) = z \times x + z \times y. \quad (2.6)$$

Commutativity: For x, y ∈ F,

$$x + y = y + x, \quad (2.7)$$

$$x \times y = y \times x. \quad (2.8)$$

Additive and multiplicative inverses: For x ∈ F there is a unique element −x ∈ F (the additive inverse of x) for which

$$x + (-x) = 0 = (-x) + x \quad (2.9)$$

For each non-zero y ∈ F there is a unique element $(\frac{1}{y})$ ∈ F (the multiplicative inverse of y) for which

$$y \times (\frac{1}{y}) = 1 = \frac{1}{y} \times y \quad (2.10)$$

- **Remarks 2.1.**

• Usually *xy* is written instead of *x* × *y*, and then we always have *xy* = *yx*.

• Because of commutativity, an above portion standards or rules are repetitive as in the sense that they are results of others (Kolman, 1996).

• When working with vectors we will dependably have a particular field of scalars at the top of the priority list and will make utilization of these guidelines.

- **Definition 2.1**

A real vector space is a set V of elements on which we have two operations ⊕ and ⊖ defined with the following properties:

(a) if u and v are any elements in V. then u ⊕ J v is in V, (We say that V is closed under the operation ⊕).
(1) u ⊕ v = v ⊕ u for all u,v in V.
(2) u ⊕ (v ⊕ w) = (n ⊕ v) ⊕ w for all u, v, w in V.
(3) There exists an element - u in V such that **u ⊕u = - u ⊕ u = 0.**
(4) If u is any element in V and c is any real number, then c ⊙ n is in V (i.e., V is closed under the operation ⊙).
(b) If u is any element in V and c is any real number, then c⊙n is in V (i.e., V is closed under the operation ⊙).
(5) c ⊙ (u ⊕ v) = c ⊙ u⊕ c ⊙ v for any u, v in V and any real number c.
(6) (c + d) ⊙ u = c ⊙ u ⊕ d ⊙ u for any u in V and any real numbers f and d.
(7) c ⊙ (l ⊙ u) = (cd) ⊙ u for any u in V and any real numbers c and d.
(8) I ⊙ u = u for any u in V.

The elements of V are called vectors: the elements of the set of real numbers R are called scalars. The operation ⊕ is called vector addition: the operation ⊙ is called scalar multiplication. The vector 0 in property (3) is called a zero vector, The vector - u in property (4) is called a negative of u.

- **Definition 2.2**

Let V be a vector space and W a nonempty subset of V. If W is a vector space with respect to the operations in V, then W is called a subspace of V.

It follows from Definition 2.2 that to verify that a subset W of a vector space V is a subspace, one must check that (a), (b), and (1) through (8) of Definition 2.1 hold. However, the next theorem says that it is enough to merely check that (a) an (b) hold to verify that a subset W of a vector space V is a subspace. Property (a) is called the closure property for ⊕, and property (b) is called the closure property for ⊙.

- **Theorem 2.1**

Let V be a vector space with operations ⊕ and ⊙ and let W be a nonempty subset of V. Then W is a subspace of V if and only if the following conditions hold:

(a) If u and v are any vectors in W, then u⊕v is in W.

(b) If c is any real number and u is any vector in W. then e ⊙ u is in W.

- **Proof**

If W is a subspace of V, then it is a vector space and (a) and (b) of Definition 4.4 hold; these are precisely (a) and (b) of the theorem

Conversely, suppose that (a) and (b) hold. We wish to show that W is a subspace of V. First, from (b) we have that (- 1) ⊙ u is in W for any u in W. From (a) we have that u ⊕ (-1) ⊙ u is in W. But u ⊕ (-1) ⊙ u = 0, so 0 is in W. Then u ⊕ 0 = u for any u in W. Finally, properties (1), (2), (5), (6), (7), and (8) hold in W because they hold in V. Hence W is a subspace of V.

- **Example 2.1**

Let W be the set of all vectors in R3 of the form $\begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \\ \boldsymbol{a+b} \end{bmatrix}$ where a and b are any real numbers. To verify Theorem 2.1 (a) and (b), we let

$$\mathbf{u} = \begin{bmatrix} a_1 \\ b_1 \\ a_1 + b_1 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} a_2 \\ b_2 \\ a_2 + b_2 \end{bmatrix}$$

be two vectors in W. Then

$$\mathbf{u} \oplus \mathbf{v} = \begin{bmatrix} a_1 + a_2 \\ b_1 + b_2 \\ (a_1 + b_1) + (a_2 + b_2) \end{bmatrix} = \begin{bmatrix} a_1 + a_2 \\ b_1 + b_2 \\ (a_1 + a_2) + (b_1 + b_2) \end{bmatrix}$$

is in W. for W consists of all those vectors whose third entry is the sum of the first two entries. Similarly,

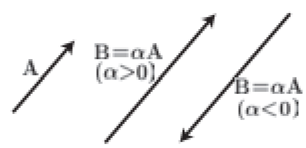
$$c \odot \begin{bmatrix} a_1 \\ b_1 \\ a_1 + b_1 \end{bmatrix} = \begin{bmatrix} ca_1 \\ cb_1 \\ c(a_1 + b_1) \end{bmatrix} = \begin{bmatrix} ca_1 \\ cb_1 \\ ca_1 + cb_1 \end{bmatrix}$$

is in W. Hence W is a subspace of R^3.

2.3 Vector Algebra

Here, we introduce a few useful operations which are defined for free vectors. Multiplication by a scalar If we multiply a vector A by a scalar α, the result is a vector $\boldsymbol{B = \alpha A}$, which has magnitude $B = |\alpha| A$. The vector $\boldsymbol{B}$, is parallel to $\boldsymbol{A}$ and points in the same direction if $\alpha > 0$. For $\alpha < 0$, the vector $\boldsymbol{B}$ is parallel to $\boldsymbol{A}$ but points in the opposite direction (antiparallel).

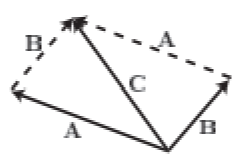
Multiplication by a scalar If we multiply a vector A by a scalar α, the result is a vector $\boldsymbol{B = \alpha A}$, which has magnitude $B = |\alpha|A$. The vector $\boldsymbol{B}$, is parallel to $\boldsymbol{A}$ and points in the same direction if $\alpha > 0$. For $\alpha < 0$, the vector $\boldsymbol{B}$ is parallel to $\boldsymbol{A}$ but points in the opposite direction (antiparallel) (Kolman, 1996).



Once we multiply an arbitrary vector, $\boldsymbol{A}$, by the inverse of its magnitude, $(1/A)$, we obtain a unit vector which is parallel to $\boldsymbol{A}$. There exist several common notations to denote a unit vector, e.g. $A\hat{}$, eA, etc. Thus, we have that $A\hat{} = A/A = A/|A|$, and $A = A\,A\hat{}$, $|A\hat{}| = 1$.

- Vector addition

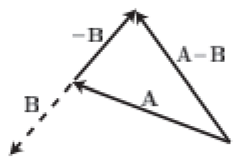
Vector addition has a very simple geometrical interpretation. To add vector $\boldsymbol{B}$ to vector $\boldsymbol{A}$, we simply place the tail of $\boldsymbol{B}$ at the head of $\boldsymbol{A}$. The sum is a vector $\boldsymbol{C}$ from the tail of $\boldsymbol{A}$ to the head of $\boldsymbol{B}$. Thus, we write $C = A + B$. The same result is obtained if the roles of $\boldsymbol{A}$ are reversed $\boldsymbol{B}$. That is, $C = A + B = B + A$. This commutative property is illustrated below with the parallelogram construction.



Since the result of adding two vectors is also a vector, we can consider the sum of multiple vectors. It can easily be verified that vector sum has the property of association, that is,

$$(A + B) + C = A + (B + C) \tag{2. 11}$$

Vector subtraction Since $\boldsymbol{A - B = A + (-B)}$, in order to subtract $\boldsymbol{B}$ from $\boldsymbol{A}$, we simply multiply $\boldsymbol{B}$ by −1 and then add (Golan, 1995).

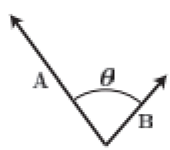


- Scalar product ("Dot" product)

This product involves two vectors and results in a scalar quantity. The scalar product between two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$, is denoted by $\boldsymbol{A \cdot B}$, and is defined as

$$A \cdot B = AB\cos\theta\,. \tag{2.12}$$

Here $\boldsymbol{\theta}$, is the angle between the vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ when they are drawn with a common origin

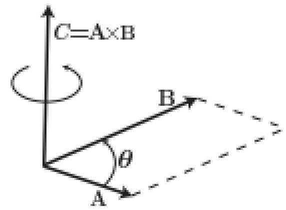


- Vector product ("Cross" product)

This product operation involves two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$, and results in a new vector $\boldsymbol{C = A \times B}$. The magnitude of $\boldsymbol{C}$ is given by,

$$C = AB\sin\theta\,, \tag{2.13}$$

where $\boldsymbol{\theta}$ is the angle between the vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ when drawn with a common origin. To eliminate ambiguity, between the two possible choices, $\boldsymbol{\theta}$ is always taken as the angle smaller than π. We can easily show that $\boldsymbol{C}$ is equal to the area enclosed by the parallelogram defined by $\boldsymbol{A}$ and $\boldsymbol{B}$. The vector $\boldsymbol{C}$ is orthogonal to both $\boldsymbol{A}$ and $\boldsymbol{B}$, i.e. it is orthogonal to the plane defined by $\boldsymbol{A}$ and $\boldsymbol{B}$. The direction of $\boldsymbol{C}$ is determined by the right-hand rule as shown (Kolman, 1996).



From this definition, it follows that

$$B \times A = -A \times B, \tag{2.14}$$

which indicates that vector multiplication is not commutative (but anticommutative). We also note that if $\boldsymbol{A} \times \boldsymbol{B} = \boldsymbol{0}$, then, either $\boldsymbol{A}$ and/or $\boldsymbol{B}$ are zero, or, $\boldsymbol{A}$ and $\boldsymbol{B}$ are parallel, although not necessarily pointing in the same direction. Thus, we also have $\boldsymbol{A} \times \boldsymbol{A} = \boldsymbol{0}$. Having defined vector multiplication, it would appear natural to define vector division. In particular, we could say that "$\boldsymbol{A}$ divided by $\boldsymbol{B}$", is a vector $\boldsymbol{C}$ such that $\boldsymbol{A} = \boldsymbol{B} \times \boldsymbol{C}$. We see immediately that there are a number of difficulties with this definition. In particular, if $\boldsymbol{A}$ is not perpendicular to $\boldsymbol{B}$, the vector $\boldsymbol{C}$ does not exist. Moreover, if $\boldsymbol{A}$ is perpendicular to $\boldsymbol{B}$ then, there are an infinite number of vectors that satisfy $\boldsymbol{A} = \boldsymbol{B} \times \boldsymbol{C}$. To see that, let us assume that $\boldsymbol{C}$ satisfies, $\boldsymbol{A} = \boldsymbol{B} \times \boldsymbol{C}$. Then, any vector $\boldsymbol{D} = \boldsymbol{C} + \beta \boldsymbol{B}$, for 3 any scalar β, also satisfies $\boldsymbol{A} = \boldsymbol{B} \times \boldsymbol{D}$, since $\boldsymbol{B} \times \boldsymbol{D} = \boldsymbol{B} \times (\boldsymbol{C} + \beta \boldsymbol{B}) = \boldsymbol{B} \times \boldsymbol{C} = \boldsymbol{A}$. We conclude therefore, that vector division is not a well-defined operation (Golan, 2007).

2.4 Summary

This chapter presented a brief review of the linear algebra as a general topic. Moreover, a review of scalars and vectors including their properties and transformations was presented.

CHAPTER THREE
SYSTEM OF LINEAR EQUATIONS AND MATRICES

This chapter introduces the basic properties of determinants and some of their applications as well as the systems of linear equations.

3.1 Systems of Linear Equations: An Introduction

To discover the break-even point and the equilibrium point we need to understand two simultaneous linear equations all together. These are two illustrations of real issues that require the solution of an arrangement of linear mathematical equations in two or more variables. In this part we take up a more orderly investigation of such frameworks. We start by considering an arrangement of two direct mathematical equations in two variables. Review that such a framework may be composed in the general structure (Gerald and Dianne, 2004).

$$ax + by = h \tag{3.1}$$

$$\mathrm{cx} + \mathrm{dy} = \mathrm{k} \tag{3.2}$$

Where a, b, c, d, h, and k are real constants and neither a and b nor c and d are both zero. Presently let's concentrate on the way of the solution of linear mathematical equations in more detail. Note that the diagram of every comparison in System (1) is a straight line in the plane, so that geometrically the answer for the system is the point(s) of intersection of the two straight lines L1 and L2. Given two lines L1 and L2, one and one and only of the next may happen:

a. L1 and L2 meet at precisely one point.

b. L1 and L2 are parallel and coincident.

c. L1 and L2 are parallel and distinct.

In the first case of figure 3, the system has a unique solution comparing to the single purpose of crossing point of the two lines. In the second case, the framework has boundlessly numerous solutions comparing to the focuses lying on the same line. At long last, in the third case, the system has no solutions on the grounds that the two lines don't meet (Howard, 2005).

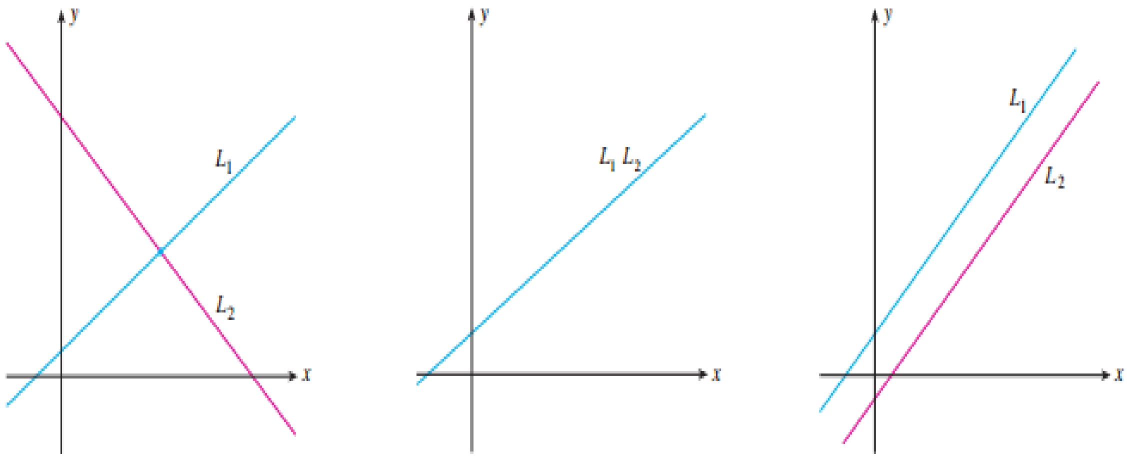


Figure 1: Different system solutions

- **Example 3.1**

Consider a system of equations with exactly one solution

$$2x - y = 1 \quad (3.3)$$

$$3x + 2y = 12 \quad (3.4)$$

If we solve the first equation for y in terms of x, we get the equation

$$y = 2x - 1 \quad (3.5)$$

Now substitute this equation for y into the second equation gives

$$3x + 2(2x - 1)2 = 12$$

$$3x + 4x - 2 = 12$$

$$7x = 14$$

$$x = 2$$

Finally, we can obtain the following by substituting this value of x into the expression for y

$$y = 2(2) - 1 = 3 \quad (3.6)$$

NOTE The result can be checked by substituting the values x = 2 and y = 3 into the equations. Thus,

$$2(2) - (3) = 3$$

$$3(2) + 2(3) = 12$$

By this verification, we can conclude that point (2, 3) lies on both lines (David, 2005).

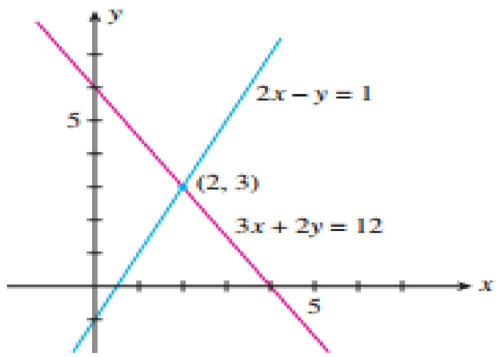


Figure 2: A system of equations with one solution

- **Example 3.2**

Consider a system of equations with infinitely many solutions

$$2x - y = 1 \quad (3.7)$$

$$6x + 3y = 3 \quad (3.8)$$

If we solve the first equation for y in terms of x, we get the equation below

$$y = 2x - 1 \quad (3.9)$$

Now let's Substitute this expression for y into the second equation

$$6x - 3(2x - 1)2 = 3$$

$$6x - 6x + 3 = 3$$

$$0 = 0$$

This is a true proclamation. This outcome takes after from the way that the second equation is proportionate to the first. Our calculations have uncovered that the solution of two mathematical equations is equal to the single mathematical equation $2x - y = 1$. In this way, any requested pair of numbers (x, y) fulfilling the mathematical equation $2x - y = 1$ (or $y = 2x - 1$) constitutes an answer for the system (Bernard and David, 2007).

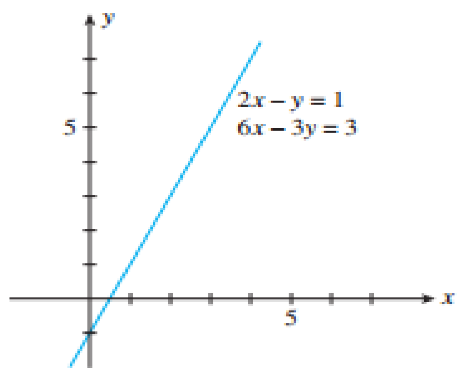


Figure 3: A system of equations with infinitely many solutions; each point on the line is a solution

- **Example 3.3**

Consider a system of equations that has no solution

$$2x - y = 1 \quad (3.10)$$

$$6x - 3y = 12 \quad (3.11)$$

The first equation is equivalent to $y = 2x - 1$. Therefore, if we substitute y into the second equation yields

$$6x - 3(2x - 1)2 = 12$$

$$6x - 6x + 3 = 12$$

$$0 = 9$$

which is plainly illogical. In this manner, there is no answer for the system of mathematical equations (Stephen et al., 2002).

To decipher this circumstance geometrically, cast both equations in the slope-intercept form, getting

$$y = 2x - 1 \quad (3.12)$$

$$y = 2x - 4 \quad (3.13)$$

We note that without a moment's delay the lines that represent these equations are parallel (each has slope 2) and distinct since the first has y-intercept -1 and the second has y-intercept -4 (Fig. 4). Systems without any solutions, for example, this one, are said to be inconsistent.

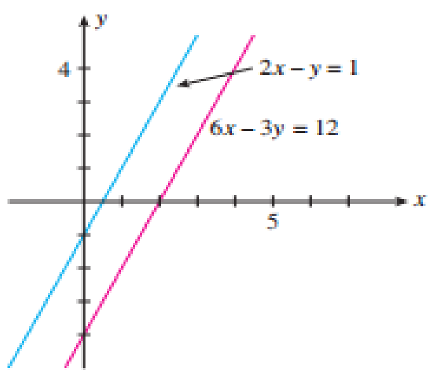


Figure 4: A system of equations with no solution

3.2 Matrices and Elementary Row Operations

In the previous we saw that changing over a linear system to an equivalent triangular system gives a calculation to illuminating the straight system. The calculation can be streamlined by acquainting matrices which represent linear systems (David, 2005).

3.2.1 what is a matrix

- **Definition 3.1**

An m × n matrix is an array of numbers with m rows and n column

For example consider this array is a 3 × 4 matrix

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

When solving a linear system by the elimination method, only the coefficients of the variables and the constants on the right-hand side are needed to find the solution. The variables are placeholders. Utilizing the structure of a matrix, we can record the coefficients and the constants by using the columns as placeholders for the variables.

$$\begin{cases} -4x_1 + 2x_2 \qquad\quad - 3x_4 = \ \ 11 \\ \ \ 2x_1 - \ \ x_2 - 4x_3 + 2x_4 = -\ 3 \\ \qquad\quad 3x_2 \qquad\quad - \ \ x_4 = \ \ \ 0 \\ -2x_1 \qquad\qquad\qquad + \ \ x_4 = \ \ \ 4 \end{cases}$$

For example, the coefficients and constants of the linear system can be recorded in matrix form as

$$\left[\begin{array}{cccc|c} -4 & 2 & 0 & -3 & 11 \\ 2 & -1 & -4 & 2 & -3 \\ 0 & 3 & 0 & -1 & 0 \\ -2 & 0 & 0 & 1 & 4 \end{array}\right]$$

This matrix is called the augmented matrix of the linear system. Notice that for an m × n linear system the augmented matrix is m × (n + 1). The augmented matrix with the last column deleted

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

is called the coefficient matrix. Notice that we always use a 0 to record any missing terms. The method of elimination on a linear system is equivalent to performing similar operations on the rows of the corresponding augmented matrix. The relationship is illustrated below:

Linear system	Corresponding augmented matrix
$\begin{cases} x + y - z = 1 \\ 2x - y + z = -1 \\ -x - y + 3z = 2 \end{cases}$	$\left[\begin{array}{rrr\|r} 1 & 1 & -1 & 1 \\ 2 & -1 & 1 & -1 \\ -1 & -1 & 3 & 2 \end{array}\right]$
Using the operations $-2E_1 + E_2 \to E_2$ and $E_1 + E_3 \to E_3$, we obtain the equivalent triangular system	Using the operations $-2R_1 + R_2 \to R_2$ and $R_1 + R_3 \to R_3$, we obtain the equivalent augmented matrix
$\begin{cases} x + y - z = 1 \\ -3y + 3z = -3 \\ 2z = 3 \end{cases}$	$\left[\begin{array}{rrr\|r} 1 & 1 & -1 & 1 \\ 0 & -3 & 3 & -3 \\ 0 & 0 & 2 & 3 \end{array}\right]$

The notation used to describe the operations on an augmented matrix is similar to the notation we introduced for equations. In the example above,

$$-2R1 + R2 \rightarrow R2$$

means replace row 2 with −2 times row 1 plus row 2. Analogous to the triangular form of a linear system, a matrix is in triangular form provided that the first nonzero entry for each row of the matrix is to the right of the first nonzero entry in the row above it.

- **Theorem 3.1**

Any one of the following operations performed on the augmented matrix, corresponding to a linear system, produces an augmented matrix corresponding to an equivalent linear system (Roger and Charles, 1990).

1. Interchanging any two rows.

2. Multiplying any row by a nonzero constant.

3. Adding a multiple of one row to another.

3.3 Solving Linear Systems with Augmented Matrices

The operations in Theorem 3.1 are called row operations. An m × n matrix A is called row equivalent to an m × n matrix B if B can be obtained from A by a sequence of row operations. The following steps summarize a process for solving a linear system (Howard, 2005).

1. Write the augmented matrix of the linear system.

2. Use row operations to reduce the augmented matrix to triangular form.

3. Interpret the final matrix as a linear system (which is equivalent to the original).

4. Use back substitution to write the solution.

Example 3.2 illustrates how we can carry out steps 3 and 4.

- **Example 3.4**

Write the augmented matrix and solve the linear system (Larry, 1998).

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \quad \textbf{b.} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right] \quad \textbf{c.} \left[\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 1 \\ 0 & 3 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

a. Reading directly from the augmented matrix, we have $x_3 = 3$, $x_2 = 2$, and $x_1 = 1$. So the system is consistent and has a unique solution.

b. In this case the solution to the linear system is $x_4 = 3$, $x_2 = 1 + x_3$, and $x_1 = 5$. So the variable x_3 is free, and the general solution is $S = \{(5, 1 + t, t, 3) \,/\, t \in \mathbb{R}\}$

c. The augmented matrix is equivalent to the linear system

$$\begin{cases} x_1 + 2x_2 + x_3 - x_4 = 1 \\ \quad\quad 3x_2 - x_3 = 1 \end{cases}$$

d. Using back substitution, we have

$$x_2 = \frac{1}{3}(1 + x_3) \quad \text{and} \quad x_1 = 1 - 2x_2 - x_{31} + x_{14} = \frac{1}{3} - \frac{5}{3}x_3 + x_{14}$$

- **Theorem 3.2**

Properties of Matrix Addition and Scalar Multiplication Let A, B, and C be m × n matrices and c and d be real numbers.

1. $A + B = B + A$

2. $A + (B + C) = (A + B) + C$

3. $c(A + B) = cA + cB$

4. $(c + d)A = cA + dA$

5. $c(dA) = (cd)A$

6. The m × n matrix with all zero entries, denoted by 0, is such that $A + 0 = 0 + A = A$.

7. For any matrix A, the matrix −A, whose components are the negative of each component of A, is such that $A + (-A) = (-A) + A = 0$ (Stephen et al., 2002).

- **Proof**

  In each case it is sufficient to show that the column vectors of the two matrices agree. We will prove property 2 and leave the others as exercises. (2) Since the matrices A, B, and C have the same size, the sums (A + B) + C and A + (B + C) are defined and also have the same size. Let $\mathbf{A_i}$, $\mathbf{B_i}$, and $\mathbf{C_i}$ denote the ith column vector of A, B, and C, respectively. Then

$$(\mathbf{A}_i + \mathbf{B}_i) + \mathbf{C}_i = \left(\begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix} + \begin{bmatrix} b_{1i} \\ \vdots \\ b_{mi} \end{bmatrix} \right) + \begin{bmatrix} c_{1i} \\ \vdots \\ c_{mi} \end{bmatrix}$$
$$= \begin{bmatrix} a_{1i} + b_{1i} \\ \vdots \\ a_{mi} + b_{mi} \end{bmatrix} + \begin{bmatrix} c_{1i} \\ \vdots \\ c_{mi} \end{bmatrix} = \begin{bmatrix} (a_{1i} + b_{1i}) + c_{1i} \\ \vdots \\ (a_{mi} + b_{mi}) + c_{mi} \end{bmatrix}$$

Since the components are real numbers, where the associative property of addition holds, we have

$$(\mathbf{A}_i + \mathbf{B}_i) + \mathbf{C}_i = \begin{bmatrix} (a_{1i} + b_{1i}) + c_{1i} \\ \vdots \\ (a_{mi} + b_{mi}) + c_{mi} \end{bmatrix}$$
$$= \begin{bmatrix} a_{1i} + (b_{1i} + c_{1i}) \\ \vdots \\ a_{mi} + (b_{mi} + c_{mi}) \end{bmatrix} = \mathbf{A}_i + (\mathbf{B}_i + \mathbf{C}_i)$$

As this holds for every column vector, the matrices (A + B) + C and A + (B + C) are equal, and we have (A + B) + C = A + (B + C).

3.4 Matrix Multiplication

Let A be an $m \times n$ matrix and B an $n \times p$ matrix; then the product AB is an $m \times p$ matrix. The ij term of AB is the dot product of the i_{th} row vector of A with the j_{th} column vector of B, so that

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

It is important to recognize that not all properties of real numbers carry over to properties of matrices. Because matrix multiplication is only defined when the number of columns of the matrix on the left equals the number of rows of the matrix on the right, it is possible for AB to exist with BA being undefined (Tomas, 2006). For example,

$$AB = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 1 & -3 \end{bmatrix} \begin{bmatrix} 3 & -2 & 5 \\ -1 & 4 & -2 \\ 1 & 0 & 3 \end{bmatrix}$$

is defined, but

$$BA = \begin{bmatrix} 3 & -2 & 5 \\ -1 & 4 & -2 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ 2 & 1 & -3 \end{bmatrix}$$

is not. As a result, we cannot interchange the order when multiplying two matrices unless we know beforehand that the matrices commute. We say two matrices A and B commute when AB = BA

3.5 Matrix Transpose

The transpose of a matrix is obtained by interchanging the rows and columns of a matrix.

- **Definition 3.2**

The transpose of a matrix is a new matrix whose rows are the columns of the original. (This makes the columns of the new matrix the rows of the original). Here is a matrix and its transpose:

$$\begin{pmatrix} 5 & 4 & 3 \\ 4 & 0 & 4 \\ 7 & 10 & 3 \end{pmatrix}^T = \begin{pmatrix} 5 & 4 & 7 \\ 4 & 0 & 10 \\ 3 & 4 & 3 \end{pmatrix}$$

The superscript "T" means "transpose".

- **Definition 3.3**

A matrix A with real entries is called symmetric if $A^T = A$.

3.6 Diagonal Matrix

An n x n matrix A = *[a_{ij}]* is called a diagonal matrix if $a_{ij} = 0$ for $i \neq j$. Thus, for a **diagonal matrix,** the terms *off* the main diagonal are all zero. Note that *O* is a diagonal matrix. A scalar matrix is a diagonal matrix whose diagonal elements are equal. The **scalar matrix** I_n= *[d_{ij}]*, where $d_{ii} = 1$ and $d_{ij} = 0$ for $i \neq j$, is called the n x n identity matrix.

- **Definition 3.4**

A n x n matrix is called **nonsingular** if there exists an n x n matrix B such that $AB = BA = I_n$; such a B is called an inverse of A. Otherwise, A is called **singular, or noninvertible.**

- **Definitions 3.5**

Let A = *[a_{ij}]* be an n x n matrix. **The determinant function**, denoted by ***det***, is defined by

$$\det(A) = \sum (\pm) a_{1j_1} a_{2j_2} \cdots a_{nj_n},$$

where the summation is over all permutations $j_1 j_2 \cdots j_n$ of the set S = (1, 2,…….,n). The sign is taken as + or - according to whether the permutation $j_1 j_2 \cdots j_n$ is even or odd.

In each term (±) $a1_{j1}$, $a2_{j2}$,….an_{jn} of **det(A),** the row subscripts are in natural order and the column subscripts are in the order $j_1 j_2 \cdots j_n$. Thus each term in **det(A),** with its appropriate sign, is a product of n entries of A with exactly one entry from each row and exactly one entry from each column. Since we slim over all permutations of S, **det(A)** has n! terms in the sum. Another notation for **det(A) is |A|.** We shall use both **det(A) and |A|.**

- **Example 3.5**

If

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

then to obtain det(A), we write down the terms $a_{1\_}a_{2\_}$ and replace the dashes with all possible elements of S2: The subscripts become 12 and 21. Now 12 is an even permutation and 21 is an odd permutation. Thus

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Hence we see that det(A) can be obtained by forming the product of the entries on the line from left to right and subtracting from this number the product of the entries on the line from right to left.

Thus, if A= $\begin{bmatrix} 2 & -3 \\ 4 & 5 \end{bmatrix}$, then |A|= (2)(5)_(-3)(4) = 22.

CHAPTER 4
LINEAR COMBINATIONS AND LINEAR INDEPENDENCE

This chapter presents an explanation of the linear combinations as well as linear independence.

4.1 Linear Combinations

For the most part, mathematics, you say that a linear combination of things is an entirety of products of those things (Poole, 2010). Along these lines, for instance, one linear combination of the functions f(x), g(x), and h(x) is

$$2f(x) + 3g(x) - 4h(x) \quad (4.1)$$

- **Definition 4.1**

A linear combination of vectors V_1, V_2, . . . , V_k in a vector space V is an expression of the form

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k \quad (4.2)$$

where the c_i's are scalars, that is, it's a whole of scalar products of them (Larry, 1998).

4.1.1 A basis for a vector space.

Some bases for vector spaces officially are known, regardless of the possibility that we haven't known them by that name. For example, in $\mathbb{R}^3$ the three vectors i = (1, 0, 0) which focuses along the x-axis, j = (0, 1, 0) which focuses along the y-axis, and k = (0, 0, 1) which focuses along the z-axis together from the standard premise for $\mathbb{R}^3$. Each vector (x, y, z) in $\mathbb{R}^3$ is an extraordinary linear combination of the standard basis vectors (Henry, 2008).

$$(x, y, z) = x_i + y_j + z_k. \quad (4.3)$$

That's the one and only linear combination of i, j, and k that gives (x, y, z).

- **Definition 4.2**

A (ordered) subset of a vector space V is a (requested) premise of V if every vector v in V may be interestingly represented as a linear combination of vectors from β.

$$v = v_1 b_1 + v_2 b_2 + \cdots + v_n b_n. \tag{4.4}$$

For a requested basis, the coefficients in that linear combination are known as the coordinates of the vector as for β.

Later on, when we study arranges in more detail, we'll compose the coordinates of a vector v as a segment vector and give it a special notation.

$$[V]\beta = \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ . \\ v_n \end{bmatrix} \tag{4.5}$$

Although we have a standard basis for $\mathbf{R_n}$, there are other bases (Lloyd and David, 1997).

- **Example 4.1**

In R^3 let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \text{and} \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

The vector

$$\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}$$

is a linear combination of V_I, V_2, and V_3 if we can find real numbers a_1, a_2, and a_3 so that

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 = \mathbf{v}. \tag{4.6}$$

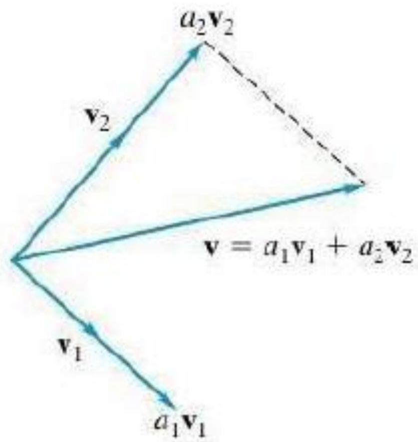


Figure 5: Linear combination of vectors

Substituting for v, V_1, V_2, and V_3, we have

$$a_1\begin{bmatrix}1\\2\\1\end{bmatrix} + a_2\begin{bmatrix}1\\0\\2\end{bmatrix} + a_3\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}2\\1\\5\end{bmatrix}.$$

Equating corresponding entries leads to the linear system (verify)

$$\begin{aligned} a_1 + a_2 + a_3 &= 2 \\ 2a_1 + a_3 &= 1 \\ a_1 + 2a_2 &= 5. \end{aligned}$$

Solving this linear system by the methods of Chapter 2 gives (verify) $a_1 = 1$, $a_2 = 2$, and $a_3 = -1$, which means that V is a linear combination of V_I, V_2, and V_3. Thus

$$\mathbf{v} = \mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3. \tag{4.7}$$

The Figure below shows V as a linear combination of V_1, V_2, and V_3.

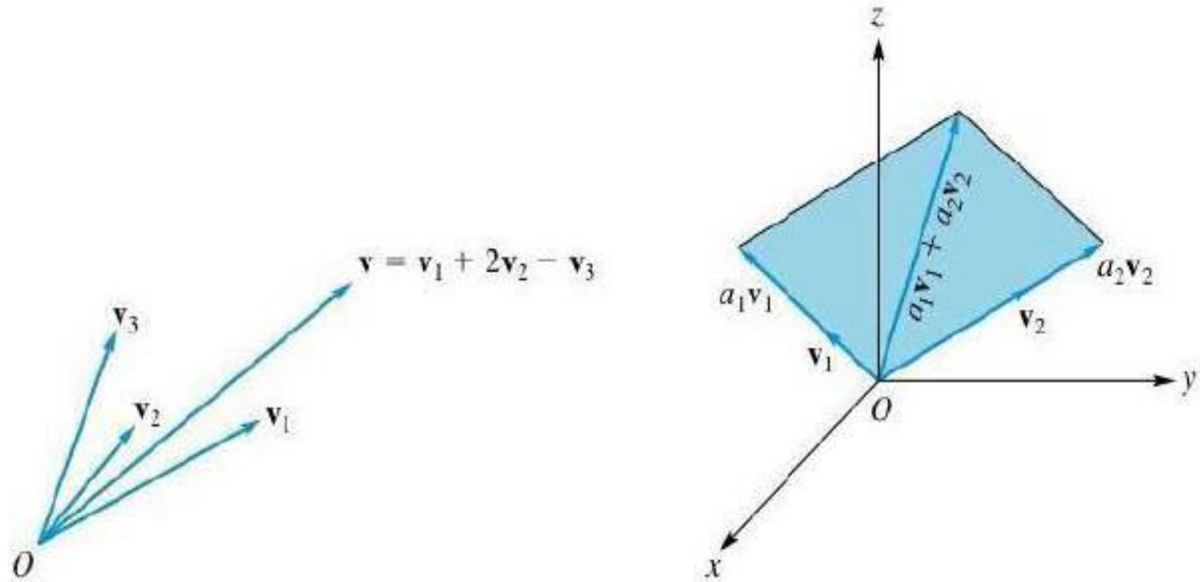


Figure 6: Linear combination of V_1, V_2, and V_3

- **Definition 4.3**

The vectors V_1, V_2 ……. V_t in a vector space V are said to be linearly dependent if there exist constants a_1, a_2, ……at, not all zero, such that

$$\sum_{j=1}^{k} a_j \mathbf{v}_j = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k = 0. \tag{4.8}$$

Otherwise, V_1, V_2 ….,V_k are called linearly independent. That is, V_1, V_2 ,…..,V_k are linearly independent if, whenever $a_1V_1 + a_2V_2 + ... + a_kV_k = 0$,

a1 = a2 =……. = ak = 0.

If S = {V_1, V_2,......,V_d},then we also say that the set S is linearly dependent or linearly independent if the vectors have the corresponding property.

- **Example 4.2**

Determine whether the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$$

are linearly independent.

- **Solution**

Forming Equation (1),

$$a_1 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + a_3 \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

we obtain the homogeneous system (verify)

$$\begin{aligned} 3a_1 + a_2 - a_3 &= 0 \\ 2a_1 + 2a_2 + 2a_3 &= 0 \\ a_1 \qquad - a_3 &= 0. \end{aligned}$$

The corresponding augmented matrix is

$$\left[\begin{array}{rrr|r} 3 & 1 & -1 & 0 \\ 2 & 2 & 2 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right],$$

whose reduced row echelon form is (verify)

$$\left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus there is a nontrivial solution

$$\begin{bmatrix} k \\ -2k \\ k \end{bmatrix}, \quad k \neq 0 \text{ (verify)},$$

so the vectors are linearly dependent.

- **Example 4.3**

Are the vectors $V_1 = [1\ 0\ 1\ 2]$, $V_2 = [0\ 1\ 1\ 2]$, and $V_3 = [1\ 1\ 1\ 3]$ in R_4 linearly dependent or linearly independent?

- **Solution**

We form Equation (1).

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 = \mathbf{0},$$

and solve for a_1, a_2, and a_3 . The resulting homogeneous system is (verify)

$$\begin{aligned} a_1 \qquad\quad + \ a_3 &= 0 \\ a_2 + \ a_3 &= 0 \\ a_1 + \ a_2 + \ a_3 &= 0 \\ 2a_1 + 2a_2 + 3a_3 &= 0. \end{aligned}$$

The corresponding augmented matrix is (verify)

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

and its reduced row echelon form is (verify)

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

Thus the only solution is the trivial solution $a_1 = a_2 = a_3 = 0$, so the vectors are linearly independent.

4.3 Testing for Linear Dependence of Vectors

There are numerous circumstances when we may wish to know whether an arrangement of vectors is linearly independent, that is if one of the vectors is some combinations of the others.

Two vectors u and v are linearly independent if the main numbers x and y fulfilling x**u**+y**v**=0 are x=y=0. On the off chance that we let

$$\vec{u} = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{and} \quad \vec{v} = \begin{bmatrix} c \\ d \end{bmatrix} \tag{4.9}$$

then xu + y**v** = 0 is equivalent to

$$0 = x\begin{bmatrix} a \\ b \end{bmatrix} + y\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \tag{4.10}$$

In the event that u and v are linearly independent, then the main answer for this arrangement of mathematical statements is the trivial solution, x=y=0. For homogeneous systems this happens exactly when the determinant is non-zero. We have now discovered a test for figuring out if a given set of vectors is linearly independent: A set of n vectors of length n is linearly independent if the matrix with these vectors as columns has a non-zero determinant. The set is obviously dependent if the determinant is zero (steven, 2006).

CHAPTER 5
LINEAR TRANSFORMATIONS

This chapter presents a brief explanation of the linear transformations in terms of examples, definitions and theorems.

5.1 Linear Transformations

- **Definition 5.1**

A linear transformation, T:U→V, is a capacity that conveys components of the vector space U (called the domain) to the vector space V (called the codomain), and which has two extra properties

1. $\mathbf{T}(\boldsymbol{u_1} + \boldsymbol{u_2}) = \mathbf{T}(\boldsymbol{u_1}) + \mathbf{T}(\boldsymbol{u_2})$ for all **u1,u2 ∈U**
2. $\mathbf{T}(\boldsymbol{\alpha}\mathbf{u}) = \boldsymbol{\alpha}\mathbf{T}(\mathbf{u})$ for all **u∈U** and all **α∈C.**

The two characterizing conditions in the meaning of a linear transformation ought to "feel linear," whatever that implies. On the other hand, these two conditions could be taken as precisely what it intends to be linear. As each vector space property gets from vector addition and scalar multiplication, so as well, every property of a linear transformation gets from these two characterizing properties. While these conditions may be reminiscent of how we test subspaces, they truly are entirely diverse, so don't befuddle the two (Defranza and Gagliardi, 2009).

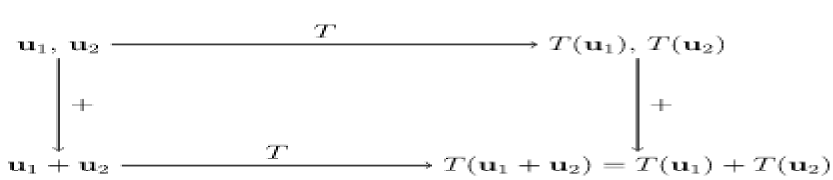


Figure 7: Definition of Linear Transformation, additive

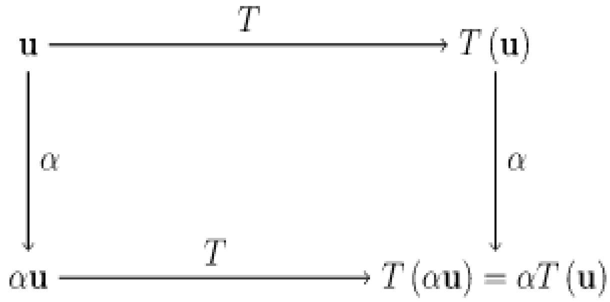


Figure 8: Definition of Linear Transformation, Multiplicative

Here are several words about notations. T is the name of the Linear Transformation, and ought to be utilized when we need to talk about the capacity in general. T(u) is the manner by which we discuss the output of the function, it is a vector in the vector space V. When we compose T(x+y)=T(x)+T(y), the plus sign on the left is the operation of vector addition in the vector space U, since x and y are components of U. The plus sign on the privilege is the operation of vector addition in the vector space V, since T(x) and T(y) are components of the vector space V. These two cases of vector addition may be uncontrollably distinctive (Gilbert, 2009).

- **Definition 5.2**

NLT: Not a linear transformation

- **Example 5.1**

Let L: $R^3 \rightarrow R^3$ be defined by

$$L\left(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}\right) = \begin{bmatrix} u_1 + 1 \\ 2u_2 \\ u_3 \end{bmatrix}.$$

To determine whether L is a linear transformation. Let

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}.$$

Then

$$L(\mathbf{u}+\mathbf{v}) = L\left(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}\right) = L\left(\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ u_3+v_3 \end{bmatrix}\right)$$
$$= \begin{bmatrix} (u_1+v_1)+1 \\ 2(u_2+v_2) \\ u_3+v_3 \end{bmatrix}.$$

On the other hand

$$L(\mathbf{u}) + L(\mathbf{v}) = \begin{bmatrix} u_1+1 \\ 2u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} v_1+1 \\ 2v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} (u_1+v_1)+2 \\ 2(u_2+v_2) \\ u_3+v_3 \end{bmatrix}.$$

Letting u1 = 1, u2 = 3, u3 = - 2, V1= 2, V2 = 4, and V3 = 1, we see that L(u + v) ≠L (u) + L (v). Hence we conclude that the function L is not a linear transformation.

- **Definition 5.3**

LTPP: Linear transformation, polynomials to polynomials

- **Example 5.2**

Let L: $P_1 \rightarrow P_2$ be defined by

$$L[p(t)] = tp(t).$$

Show that L is a linear transformation.

- **Solution**

Let p(t) and q(t) be vectors in P_1 and let c be a scalar. Then

$$\begin{aligned} L[p(t) + q(t)] &= t[p(t) + q(t)] \\ &= tp(t) + tq(t) \\ &= L[p(t)] + L[q(t)], \end{aligned}$$

And

$$\begin{aligned} L[cp(t)] &= t[cp(t)] \\ &= c[tp(t)] \\ &= cL[p(t)]. \end{aligned}$$

Hence L is a linear transformation.

5.2 Properties of Linear Transformation

Let V and W be two vector spaces. Suppose T: V → W is a linear transformation (Gilbert, 2014). Then

1. T(0) = 0.

2. $T(-v) = -T(v)$ for all $v \in V$.

3. $T(u - v) = T(u) - T(v)$ for all $u, v \in V$

4. If $v = c_1v_1 + c_2v_2 + \cdots + c_nv_n$

Then

$T(v) = T(c_1v_1 + c_2v_2 + \cdots + c_nv_n) = c_1T(v_1) + c_2T(v_2) + \cdots + c_nT(v_n)$.

- **Proof**

By property (2) we have

$$T(0) = T(00) = 0T(0) = 0. \quad (5.1)$$

So, (1) is proved. Similarly,

$$T(-v) = T((-1)v) = (-1)T(v) = -T(v). \quad (5.2)$$

So, (2) is proved. Then, by property (1) of the definition 5.1, we have

$$T(u - v) = T(u + (-1)v) = T(u) + T((-1)v) = T(u) - T(v). \quad (5.3)$$

The last equality follows from (2). So, (3) is proved. To prove (4), we use induction, on n. For n= 1: we have

$$T(c_1v_1) = c_1T(v_1). \quad (5.4)$$

For n = 2, by the two properties of definition 5.1, we have

$$T(c_1v_1 + c_2v_2) = T(c_1v_1) + T(c_2v_2) = c_1T(v_1) + c_2T(v_2). \quad (5.5)$$

So, (4) is prove for n = 2. Now, we assume that the formula (4) is valid for n − 1 vectors and proves it for n. We have

$$T(c_1v_1 + c_2v_2 + \cdot\cdot\cdot + c_nv_n) = T(c_1v_1 + c_2v_2 + \cdot\cdot\cdot + c_n{-}1v_n{-}1)+T(c_nv_n) = (c_1T(v_1) + c_2T(v_2) + \cdot\cdot\cdot + c_n{-}1T(v_n{-}1)) + c_nT(v_n). \quad (5.6)$$

So, the proof is complete.

5.4 Linear Transformations Given by Matrices

- **Theorem 5.2**

Let A be a matrix of size m × n. Given a vector

$$V = \begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{bmatrix} \in \mathbb{R}^n \quad \text{define} \quad T(v) = Av = A\begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{bmatrix}.$$

Then T is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ (Katta, 2014).

- Proof

From properties of matrix multiplication, for $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and scalar c we have

$\mathrm{T}(\mathbf{u}+\mathbf{v}) = \mathrm{A}(\mathbf{u}+\mathbf{v}) = \mathrm{A}(\mathrm{u}) + \mathrm{A}(\mathbf{v}) = \mathrm{T}(\mathbf{u}) + \mathrm{T}(\mathbf{v})$ and $\mathrm{T}(c\mathbf{u}) = \mathrm{A}(c\mathbf{u}) = c\mathrm{A}\mathbf{u} = c\mathrm{T}(\mathbf{u})$.

The proof is complete (Otto, 2004).

- **Example 5.3**

Let L: $R_2 \rightarrow R_2$ be defined by

$$L\left(\begin{bmatrix} u_1 & u_2 \end{bmatrix}\right) = \begin{bmatrix} u_1^2 & 2u_2 \end{bmatrix}.$$

Is L a linear transformation?

- **Solution**

Let

$$\mathbf{u} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}.$$

Then

$$\begin{aligned} L(\mathbf{u}+\mathbf{v}) &= L\left(\begin{bmatrix} u_1 & u_2 \end{bmatrix} + \begin{bmatrix} v_1 & v_2 \end{bmatrix}\right) \\ &= L\left(\begin{bmatrix} u_1+v_1 & u_2+v_2 \end{bmatrix}\right) \\ &= \begin{bmatrix} (u_1+v_1)^2 & 2(u_2+v_2) \end{bmatrix}. \end{aligned}$$

On the other hand

$$L(\mathbf{u}) + L(\mathbf{v}) = \begin{bmatrix} u_1^2 & 2u_2 \end{bmatrix} + \begin{bmatrix} v_1^2 & 2v_2 \end{bmatrix}$$
$$= \begin{bmatrix} u_1^2 + v_1^2 & 2(u_2 + v_2) \end{bmatrix}.$$

Since there are some choices of u and v such that **L(u + v) ≠ L(u) + L(v).** we conclude that L is not a linear transformation.

CHAPTER 6
APPLICATIONS OF EIGENVALUES AND EIGENVECTORS

This chapter presents a detailed introduction of the eigenvectors and eigenvalues. It explains the methods to find the eigenvalues and eigenvectors in a matrix. Moreover, it discusses the numerous applications of eigenvalues and eigenvectors in different fields.

6.1 Introduction to Eigenvalues and Eigenvectors

If we multiply an n x n matrix by an n x 1 vector we will get a new n x 1 vector back. In other words,

$$A\vec{\eta} = \vec{y} \tag{6.1}$$

What we want to know is if it is possible for the following to happen. Instead of just getting a brand new vector out of the multiplication is it possible instead to get the following,

$$A\vec{\eta} = \lambda\vec{\eta} \tag{6.2}$$

In other words is it possible, at least for certain λ and $\vec{\eta}$, to have matrix multiplication be the same as just multiplying the vector by a constant? Of course, we probably wouldn't be talking about this if the answer was no. So, it is possible for this to happen, however, it won't happen for just any value of λ or $\vec{\eta}$. If we do happen to have a λ and $\vec{\eta}$ for which this works (and they will always come in pairs) then we call λ an **eigenvalue** of A and $\vec{\eta}$ an **eigenvector** of A (Jolliffe, 1986).

So, how do we go about find the eigenvalues and eigenvectors for a matrix? Well first notice that if $\vec{\eta} = \vec{0}$ then (6.1) is going to be true for any value of λ and so we are going to make the assumption that $\vec{\eta} \neq \vec{0}$. With that out of the way let's rewrite (6.1) a little.

$$A\vec{\eta} - \lambda\vec{\eta} = \vec{0}$$
$$A\vec{\eta} - \lambda I_n\vec{\eta} = \vec{0}$$
$$\left(A - \lambda I_n\right)\vec{\eta} = \vec{0}$$

Notice that before we factored out the $\vec{\eta}$ we added in the appropriately sized identity matrix. This is equivalent to multiplying things by a one and so doesn't change the value of anything. We needed to do this because without it we would have had the difference of a matrix, *A*, and a constant, λ, and this can't be done. We now have the difference of two matrices of the same size which can be done (Janardan et al., 2004).

So, with this rewrite we see that

$$\left(A - \lambda I_n\right)\vec{\eta} = \vec{0} \tag{6.3}$$

is equivalent to (6.1). In order to find the eigenvectors for a matrix we will need to solve a homogeneous system. Recall the fact from the previous section that we know that we will either have exactly one solution ($\vec{\eta} = \vec{0}$) or we will have infinitely many nonzero solutions. Since we've already said that don't want $\vec{\eta} = \vec{0}$ this means that we want the second case.

Knowing this will allow us to find the eigenvalues for a matrix. We will need to determine the values of λ for which we get,

$$\det\left(A - \lambda I\right) = 0$$

Once we have the eigenvalues we can then go back and determine the eigenvectors for each eigenvalue. Let's take a look at a couple of quick facts about eigenvalues and eigenvectors (Jolliffe, 1986).

Fact

If *A* is an *n* x *n* matrix then $\det\left(A - \lambda I\right) = 0$ is an n^{th} degree polynomial. This polynomial is called the characteristic polynomial. To find eigenvalues of a matrix all we need to do is solve a polynomial. That's generally not too bad provided we keep *n* small. Likewise this fact also tells us that for an *n* x *n* matrix, *A*, we will have *n* eigenvalues if we include all repeated eigenvalues (Mashal et al., 2005).

- **Example 6.1**

Find the eigenvalues and eigenvectors of the following matrix.

$$A = \begin{pmatrix} 2 & 7 \\ -1 & -6 \end{pmatrix}$$

- **Solution**

The first thing that we need to do is find the eigenvalues. That means we need the following matrix,

$$A - \lambda I = \begin{pmatrix} 2 & 7 \\ -1 & -6 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2-\lambda & 7 \\ -1 & -6-\lambda \end{pmatrix}$$

In particular we need to determine where the determinant of this matrix is zero.

$$\det(A - \lambda I) = (2-\lambda)(-6-\lambda) + 7 = \lambda^2 + 4\lambda - 5 = (\lambda + 5)(\lambda - 1)$$

So, it looks like we will have two simple eigenvalues for this matrix, $\lambda_1 = -5$ and $\lambda_2 = 1$. We will now need to find the eigenvectors for each of these. Also note that according to the fact above, the two eigenvectors should be linearly independent (Smith, 2002).

To find the eigenvectors we simply plug in each eigenvalue into (6.2) and solve. So, let's do that.

$\lambda_1 = -5$:

In this case we need to solve the following system.

$$\begin{pmatrix} 7 & 7 \\ -1 & -1 \end{pmatrix} \vec{\eta} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Recall that officially to solve this system we use the following augmented matrix.

$$\begin{pmatrix} 7 & 7 & 0 \\ -1 & -1 & 0 \end{pmatrix} \underset{\Rightarrow}{\tfrac{1}{7}R_1 + R_2} \begin{pmatrix} 7 & 7 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Upon reducing down we see that we get a single equation

$$7\eta_1 + 7\eta_2 = 0 \qquad \Rightarrow \qquad \eta_1 = -\eta_2$$

that will yield an infinite number of solutions. This is expected behavior, so we would get infinitely many solutions.

Notice as well that we could have identified this from the original system. This won't always be the case, but in the 2 x 2 case we can see from the system that one row will be a multiple of the other and so we will get infinite solutions. From this point on we won't be actually solving systems in these cases. We will just go straight to the equation and we can use either of the two rows for this equation (Smith, 2002).

Now, let's get back to the eigenvector, since that is what we were after. In general then the eigenvector will be any vector that satisfies the following,

$$\vec{\eta} = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} -\eta_2 \\ \eta_2 \end{pmatrix} \qquad , \eta_2 \neq 0$$

To get this we used the solution to the equation that we found above.

We really don't want a general eigenvector however so we will pick a value for η_2 to get a specific eigenvector. We can choose anything (except $\eta_2 = 0$), so pick something that will make the eigenvector "nice". Note as well that since we've already assumed that the eigenvector is not zero we must choose a value that will not give us zero, which is why we want to avoid $\eta_2 = 0$ in this case. Here's the eigenvector for this eigenvalue.

$$\vec{\eta}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \qquad \text{using } \eta_2 = 1$$

Now we get to do this all over again for the second eigenvalue.

$\lambda_2 = 1$:

We'll do much less work with this part than we did with the previous part. We will need to solve the following system.

$$\begin{pmatrix} 1 & 7 \\ -1 & -7 \end{pmatrix} \vec{\eta} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Clearly both rows are multiples of each other and so we will get infinitely many solutions. We can choose to work with either row (Mashal et al., 2005). Doing this gives us,

$$\eta_1 + 7\eta_2 = 0 \qquad\qquad \eta_1 = -7\eta_2$$

Note that we can solve this for either of the two variables. The eigenvector is then,

$$\vec{\eta} = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} -7\eta_2 \\ \eta_2 \end{pmatrix} \qquad ,\eta_2 \neq 0$$

$$\vec{\eta}^{(2)} = \begin{pmatrix} -7 \\ 1 \end{pmatrix}, \qquad\qquad \text{using } \eta_2 = 1$$

Summarizing we have,

$$\lambda_1 = -5 \qquad\qquad \vec{\eta}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 1 \qquad\qquad \vec{\eta}^{(1)} = \begin{pmatrix} -7 \\ 1 \end{pmatrix}$$

Note that the two eigenvectors are linearly independent as predicted.

6.2 Applications of Eigenvectors and Eigenvalues

Many applications of matrices in both engineering and science utilize eigenvalues and, sometimes, eigenvectors. Control theory, vibration analysis, electric circuits, advanced dynamics and quantum mechanics are just a few of the application areas. Many of the applications involve the use of eigenvalues and eigenvectors in the process of transforming a given matrix into a diagonal matrix and we discuss this process in this Section. We then go on to show how this process is invaluable in solving coupled differential equations and the applications of eigenvalues and eigenvectors in Principal Components Analysis (Boldrimi et al., 1984).

Numerous applications of matrices; in both engineering and science use eigenvalues and, in some cases, eigenvectors. Control hypothesis, vibration examination, electric circuits, propelled motion and quantum mechanics are only a couple of the application zones. Large portions of the applications include the utilization of eigenvalues and eigenvectors during the time spent changing a given matrix into a diagonal matrix and we discuss this procedure in this Section.

6.2.1 Diagonalization of a matrix with distinct eigenvalues

Diagonalization means transforming a non-diagonal matrix into an equivalent matrix which is diagonal and hence is simpler to deal with. A matrix A with distinct eigenvalues has eigenvectors which are linearly independent (Boldrimi et al., 1984). If we form a matrix P whose columns are these eigenvectors, it can then be shown that

$\det(P) \neq 0$

so that P^{-1} exists.

The product $P^{-1}AP$ is then a diagonal matrix D whose diagonal elements are the eigenvalues

of A. Thus if $\lambda_1, \lambda_2, \ldots \lambda_n$ are the distinct eigenvalues of A with associated eigenvectors $X_1, X_2, \ldots, X_n$ respectively:

$$P = [X_1 : X_2 : X_3 : \ldots \ldots \ldots \ldots \ldots \ldots X_n] \qquad (6.4)$$

will produce a product

$$P^{-1}AP = D = \begin{bmatrix} \lambda_1 & 0 & \ldots\ldots..0 \\ 0 & \lambda_2 & \ldots\ldots..0 \\ & . & \\ & . & \\ 0 & 0 & \ldots\ldots..\lambda_n \end{bmatrix} \quad (6.5)$$

We see that the order of the eigenvalues in D matches the order in which P is formed from the eigenvectors.

Note 6.1

(a) The matrix P is called the modal matrix of A.

(b) Since D, as a diagonal matrix, has eigenvalues λ_1, λ_2,...,λ_n which are the same as those of A then the matrices D and A are said to be similar. The transformation of A into D using $P^{-1}AP =$ $D =$ is said to be a similarity transformation.

- **Example 6.2**

Let A = $\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$. The eigenvalues of A are $\lambda_1 = 2$ and $\lambda_2 = 3$ and associated vectors are

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

respectively. Thus

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \quad \text{(verify).}$$

Hence

$$P^{-1}AP = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

On the other hand, if we let $\lambda_1 = 3$ and $\lambda_2 = 2$, then

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix};$$

$$P = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}.$$

And

$$P^{-1}AP = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}.$$

6.2.2 Systems of linear differential equations-Real, distinct eigenvalue

Now, it is time to start solving systems of differential equations. We've seen that solutions to the system,

$$\vec{x}' = A\vec{x} \tag{6.6}$$

will be of the form

$$\vec{x} = \vec{\eta}\mathbf{e}^{\lambda t} \tag{6.7}$$

where λ and $\vec{\eta}$ are eigenvalues and eigenvectors of the matrix *A*. We will be working with 2 x 2 systems so this means that we are going to be looking for two solutions, $\vec{x}_1(t)$ and $\vec{x}_2(t)$, where the determinant of the matrix,

$$X = (\vec{x}_1 \;\; \vec{x}_2) \tag{6.8}$$

is nonzero.

We are going to start by looking at the case where our two eigenvalues, λ_1 and λ_2 are real and distinct. In other words they will be real, simple eigenvalues. Recall as well that the eigenvectors for simple eigenvalues are linearly independent. This means that the solutions we get from these will also be linearly independent (Smith, 2002). If the solutions are linearly independent the matrix *X* must be nonsingular and hence these two solutions will be a fundamental set of solutions. The general solution in this case will then be,

$$\vec{x}(t) = c_1 \mathbf{e}^{\lambda_1 t} \vec{\eta}^{(1)} + c_2 \mathbf{e}^{\lambda_2 t} \vec{\eta}^{(2)} \tag{6.9}$$

- **Example 6.3**

Solve the following IVP.

$$\vec{x}' = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \vec{x}, \qquad \vec{x}(0) = \begin{pmatrix} 0 \\ -4 \end{pmatrix}$$

- **Solution**

So, the first thing that we need to do is find the eigenvalues for the matrix.

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{vmatrix} \\ &= \lambda^2 - 3\lambda - 4 \\ &= (\lambda+1)(\lambda-4) \qquad \Rightarrow \qquad \lambda_1 = -1,\ \lambda_2 = 4 \end{aligned}$$

Now let’s find the eigenvectors for each of these.

$\lambda_1 = -1$:

We’ll need to solve,

$$\begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \Rightarrow \qquad 2\eta_1 + 2\eta_2 = 0 \qquad \Rightarrow \qquad \eta_1 = -\eta_2$$

The eigenvector in this case is,

$$\vec{\eta} = \begin{pmatrix} -\eta_2 \\ \eta_2 \end{pmatrix} \qquad \Rightarrow \qquad \vec{\eta}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \qquad \eta_2 = 1$$

$$\lambda_2 = 4 :$$

We'll need to solve,

$$\begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \Rightarrow \qquad -3\eta_1 + 2\eta_2 = 0 \quad \Rightarrow \quad \eta_1 = \frac{2}{3}\eta_2$$

The eigenvector in this case is,

$$\vec{\eta} = \begin{pmatrix} \frac{2}{3}\eta_2 \\ \eta_2 \end{pmatrix} \qquad \Rightarrow \qquad \vec{\eta}^{(2)} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad \eta_2 = 3$$

Then general solution is then,

$$\vec{x}(t) = c_1 \mathbf{e}^{-t}\begin{pmatrix} -1 \\ 1 \end{pmatrix} + c_2 \mathbf{e}^{4t}\begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Now, we need to find the constants. To do this we simply need to apply the initial conditions.

$$\begin{pmatrix} 0 \\ -4 \end{pmatrix} = \vec{x}(0) = c_1\begin{pmatrix} -1 \\ 1 \end{pmatrix} + c_2\begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

All we need to do now is multiply the constants through and we then get two equations (one for each row) that we can solve for the constants. This gives,

$$\left.\begin{aligned} -c_1 + 2c_2 &= 0 \\ c_1 + 3c_2 &= -4 \end{aligned}\right\} \qquad \Rightarrow \qquad c_1 = -\frac{8}{5},\ c_2 = -\frac{4}{5}$$

The solution is then,

$$\vec{x}(t) = -\frac{8}{5}\mathbf{e}^{-t}\begin{pmatrix} -1 \\ 1 \end{pmatrix} - \frac{4}{5}\mathbf{e}^{4t}\begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

6.2.3 PCA based eigenvectors and eigenvalues

Principal Components Analysis (PCA) is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. It is one of several statistical tools available for reducing the dimensionality of a data set based on calculating eigenvectors and eigenvalues of the input data. Since patterns in data can be hard to find in data of high dimension, where the luxury of graphical representation is not available, PCA is a powerful tool for analyzing data. The other main advantage of PCA is that once you have found these patterns in the data, and you compress the data, i.e. by reducing the number of dimensions, without much loss of information. This technique used in image compression, as we will see in a later section. This chapter will take you through the steps you needed to perform a Principal Components Analysis on a set of data (Rafael, 2012).

- **Definition 6.1**

Let X_{jk} indicate the particular value of the k^{th} variable that is observed on the j^{th} item. We let *n* be the number of items being observed and *p* the number of variables measured. Such data are organized and represented by a rectangular matrix *X* given by a multivariate data matrix.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2k} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{j1} & x_{j2} & \cdots & x_{jk} & \cdots & x_{jp} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} & \cdots & x_{np} \end{bmatrix},$$

In a single-variable case where the matrix *X* is *n* x 1,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad (6.10)$$

The mean

$$\overline{x} = \frac{1}{n}\sum_{j=1}^{n} x_j \tag{6.11}$$

And the variance

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_j - \overline{x})^2. \tag{6.12}$$

In addition, the square root of the sample variance is known as the sample standard deviation.

- **Example 6.4**

If the matrix

$$X = [97\ 92\ 90\ 87\ 85\ 83\ 83\ 78\ 72\ 71\ 70\ 65]^T$$

is the set of scores out of 100 for an exam in linear algebra, then the associated descriptive statistics arc $\overline{\boldsymbol{x}} \approx 81$, $s^2 \approx 90.4$, and the standard deviation $s \approx 9.5$.

Mean of the k^{th} variable

$$\overline{x}_k = \frac{1}{n}\sum_{j=1}^{n} x_{jk}, \quad k = 1, 2, \ldots, p. \tag{6.13}$$

Variance of the k^{th} variable

$$s_k^2 = \frac{1}{n}\sum_{j=1}^{n}(x_{jk} - \overline{x}_k)^2, \quad k = 1, 2, \ldots, p. \tag{6.14}$$

For convenience of matrix notation, we shall use the alternative notation S_{kk} for the variance of the k^{th} variable; that is,

$$s_{kk} = s_k^2 = \frac{1}{n}\sum_{j=1}^{n}(x_{jk} - \overline{x}_k)^2, \quad k = 1, 2, \ldots, p. \tag{6.15}$$

A measure of the linear association between a pair of variables is provided by the notion of covariance. The measure of association between the i^{th} and k^{th} variables in the multivariate data matrix X is given by

$$\text{covariance} = s_{ik} = \frac{1}{n}\sum_{j=1}^{n}(x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k), \quad i = 1, 2, \ldots, p, \quad k = 1, 2, \ldots, p, \tag{6.16}$$

which is the average product of the deviations from their respective means. It follows that $s_{jk} = s_{ki}$, for all *i* and *k*, and that for *i* = *k*, the covariance is just the variance, $s^2_k = .s_{kk}$

Matrix of variances and covariances =

$$S_n = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{bmatrix}. \tag{6.17}$$

The matrix S_n is a symmetric matrix whose diagonal entries arc the sample variances and the subscript *n* is a notational device to remind us that the divisor *n* was used to compute the variances and covariances. The matrix S_n is called the covariance matrix.

- **Theorem 6.1**

Let S_n be the p x p covariance matrix associated with the multivariate data matrix X. Let the eigenvalues of S_n be λ_j, j = 1, 2,….,p $\lambda_1 \geq \lambda_2 \geq \ldots\ldots \geq \lambda_p \geq 0$, and let the associated orthonormal eigenvectors be $\mathbf{u}_j$, j = 1, 2 ....., p. Then the i^{th} principal component $\mathbf{y_i}$ is given by the linear combination of the columns of **X**, where the coefficients are the entries of the eigenvector $\mathbf{u_i}$; that is,

$$\mathbf{y_i} = \boldsymbol{i^{th}} \textbf{ principal component} = \mathbf{Xu_i}$$

- **Example 6.5**

Let $$X = \begin{bmatrix} 39 & 21 \\ 59 & 28 \\ 18 & 10 \\ 21 & 13 \\ 14 & 13 \\ 22 & 10 \end{bmatrix}.$$

Find the covariance matrix of X?

- **Solution**

We find that the means are

$$\bar{x}_1 \approx 28.8 \quad \text{and} \quad \bar{x}_2 \approx 15.8,$$

and thus we take the matrix of means as

$$\bar{\mathbf{x}} = \begin{bmatrix} 28.8 \\ 15.8 \end{bmatrix}.$$

The variances are

$$s_{11} \approx 243.1 \quad \text{and} \quad s_{22} \approx 43.1,$$

While the covariances are

$$s_{12} = s_{21} \approx 97.8.$$

Hence we take the covariance matrix as

$$S_n = \begin{bmatrix} 243.1 & 97.8 \\ 97.8 & 43.1 \end{bmatrix}.$$

- **Example 6.6**

Determine the PCA of y_1 and y_2 of covariance matrix in example 6.5?

- **Solution:**

we determined the eigenvalues of the matrix S_n .

$\lambda_1 = 282.9744$ and $\lambda_2 = 3.2256$

and associated eigenvectors

$$\mathbf{u}_1 = \begin{bmatrix} 0.9260 \\ 0.3775 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} 0.3775 \\ -0.9260 \end{bmatrix}.$$

Then, using Theorem 6.1 we find that the first principal component is

$$y1= 0.9260col_1(X) + 0.3775col_2(X)$$

and the second principal component is

$$y2=0.3775col_1(X) – 0.3775col_2(X)$$

6.2.4 PCA for image compression

Principal Component Analysis – PCA was used for the recognition of patterns and compression of digital images used in Medicine. The description of Principal Component Analysis is made by means of the explanation of eigenvalues and eigenvectors of a matrix. This concept is presented on a digital image collected in the clinical routine of a hospital, based on the functional aspects of a matrix. The analysis of potential for recovery of the original image was made in terms of the rate of compression obtained.

Principal Components Analysis (PCA) is a mathematical formulation used in the reduction of data dimensions. Thus, the PCA technique allows the identification of standards in data and their expression in such a way that their similarities and differences are emphasized. Once patterns are found, they can be compressed, i.e., their dimensions can be reduced without much loss of information. In summary, the PCA formulation may be used as a digital image compression algorithm with a low level of loss (Rafael, 2012).

Use of the PCA technique in data dimension reduction is justified by the easy representation of multidimensional data, using the information contained in the data covariance matrix, principles of linear algebra and basic statistics. The studies carried out by Mashal (Mashal et a., 2005) adopted the PCA formulation in the selections of images from a multimedia database. According to Smith (smith, 2002), PCA is an authentic image compression algorithm with minimal loss of information. The relevance of this work is in the performance evaluation of the PCA formulation in compressing digital images from the measurement of the degree of compression and the degree of information loss that the PCA introduces into the compressed images in discarding some principal components.

6.2.4.1 MRI (Magnetic Resonance Imaging) image compression using PCA

The steps normally followed in a PCA of a digital image can now be established:

Step 1: In the computational model of a digital image, in equation (6.10), the variables X1, X2 ,...,Xp are the columns of the image. The PCA is begun by coding (correcting) the image to that its columns have zero means and unitary variances. This is common, in order to avoid one or the other of the columns having undue influence on the principal components (Gonzalez and Woods, 1992)

$$\text{image corrected by the mean} = \text{image} - \text{mean of the image}$$

Step 2: The covariance matrix C is calculated using equation (6.16), implemented computationally, that is:

$$\text{covImage} = \text{image corrected by the mean} \times (\text{image corrected by the mean})^T$$

Step 3: The eigenvalues l_1 ,l_2 ,...,l_p and the corresponding eigenvectors a_1 , a_2 ,..., a_p . are calculated.

Step 4: The value of a vector of characteristics is obtained, a matrix with vectors containing the list of eigenvectors (matrix columns) of the covariance matrix (6.16).

$$vc = (av1 , av2 , av3 ,..., avn)$$

Step 5: The final data are obtained, that is, a matrix with all the eigenvectors (components) of the covariance matrix.

$$\text{finaldata} = vc^T \times (\text{Image - mean})^T$$

Step 6: The original image is obtained from the final data without compression using the expression Image

$$T = (vc)^T \times \text{finaldata} + \text{mean}^T$$

Step 7: Any components that explain only a small portion of the variation in data for the effect of image compression are discarded. The eliminations have the effect of reducing the quantity of eigenvectors of the characteristics vectors and can produce final data with a smaller dimension.

- **Compression ration**

According to Castro (Castro, 2010), low-loss compression afforded by the present method may be expressed in terms of the compression factor of (r) and of the mean squared error (MSE) committed in the approximation of A (original image) by Ã (image obtained from the disposal of some of the components) (Gonzalez and Woods, 1992). The compression factor is defined by:

$$\rho = \frac{\text{Unit of memory required to represent } \tilde{A}}{\text{Unit of memorial required to represent } A} \qquad (6.18)$$

And the MSE committed in the approximation of A by Ã is:

$$MSE = \sum_{i=0}^{L-1} (\tilde{a}_i - a_i)^T (\tilde{a}_i - a_i) \qquad (6.19)$$

- **Example 6.7**

Recovering a TIFF image with 512*512 pixels with all the components (512) of image covariance matrix (without compression, i.e., steps 1 to 6).

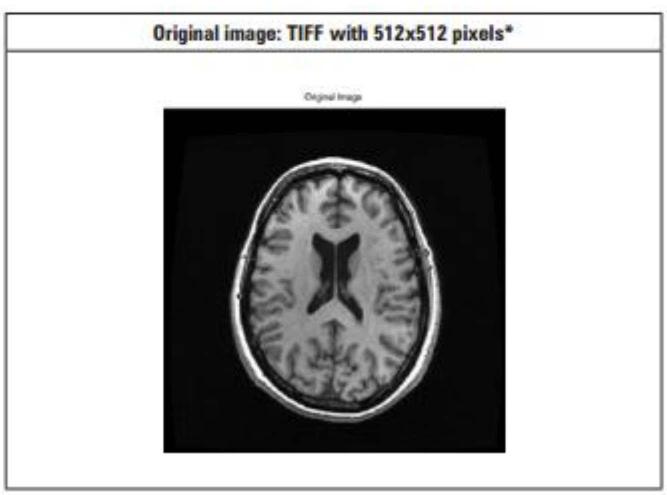


Figure 9: MRI original image (512*512) (Rafael, 2012)

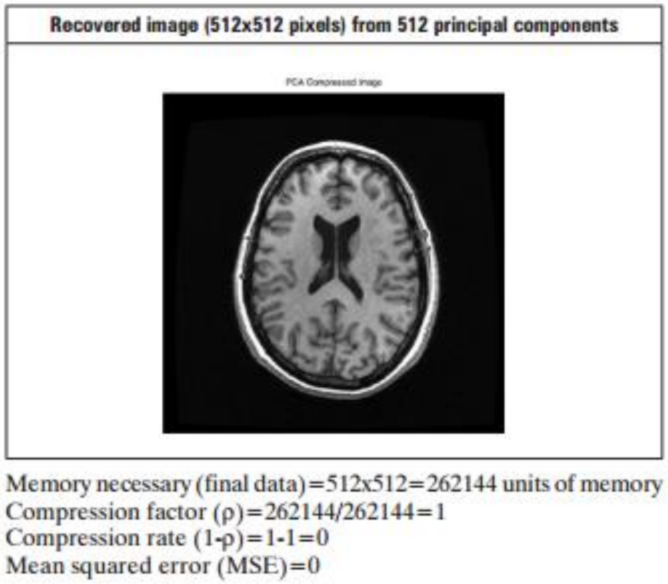


Figure 10: Recovered image without compression (Rafael, 2012)

- **Example 6.8**

Recovery of a TIFF image with 512x512 pixels with 112 principal components of the covariance matrix of the image (with compression, that is, steps from 1 to 5 to 7).

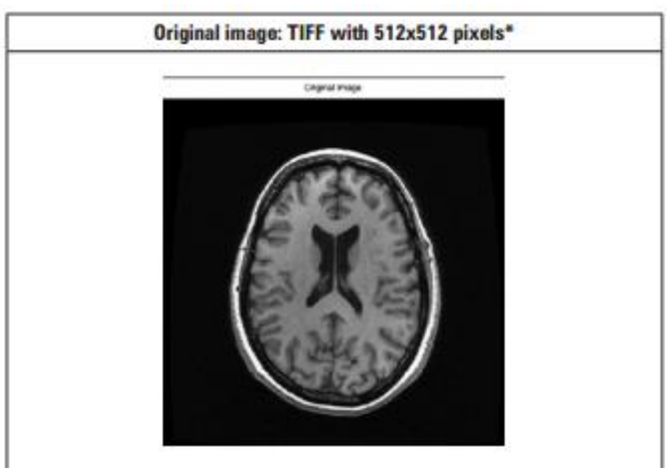
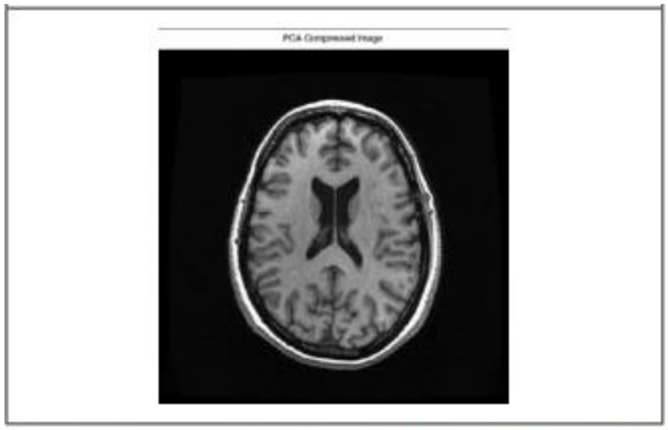


Figure 11: MRI brain original image (Rafael, 2012)



Memory necessary (final data) = 112x512 = 57344 units of memory
Compression factor (ρ) = 57344/262144 = 0.219
Compression rate (1-ρ) = 1-0.219 = 0.781
Mean squared error (MSE) = 0213

Figure 12: compressed image using PCA (Rafael, 2012)

- **Example 6.9**

Recovery of an image with 32 principal components of the image covariance matrix (with compression).

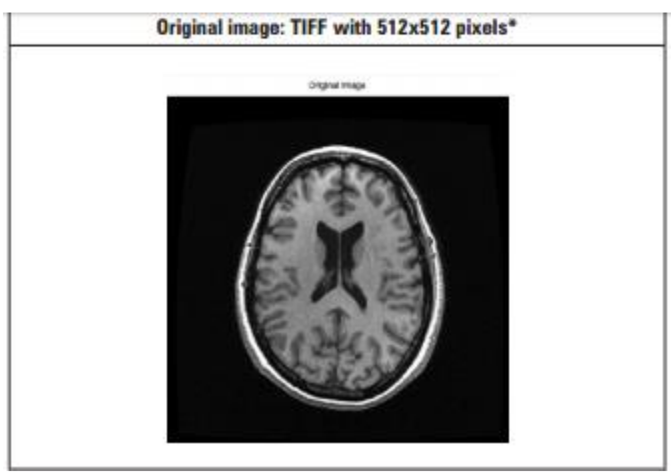


Figure 13: Original MRI image 3 (Rafael, 2012)

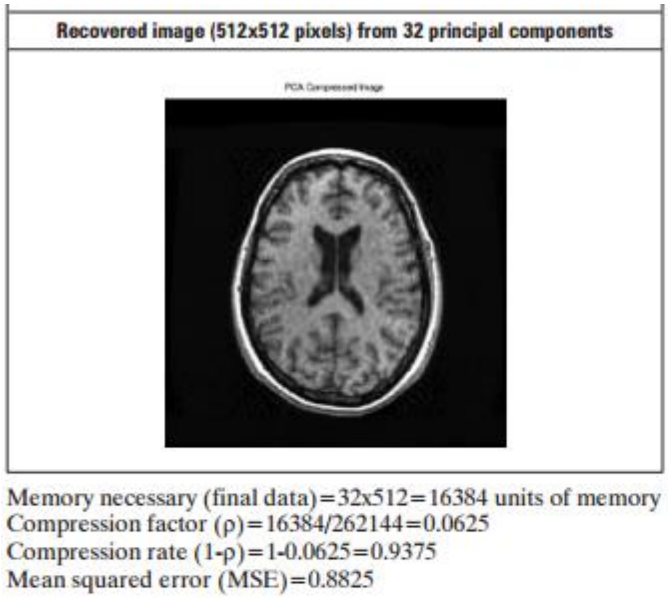


Figure 14: Compressed MRI image (32*512) (Rafael, 2012)

6.3 Results Discussion

Examples 6.3 to 6.5 show the effects of the reduction in number of principal components (elevation of the image compression rate) in the increased loss of information. This application may bring great savings in storage of medical images. However, the level of information preserved depends on the parameters (compression rate), and should be modulated by the user's interest. The higher the compression rate (the fewer principal components are used in the characteristics vector) the more degraded the quality of the image recovered (Example 6.5). In certain applications, such as brain function images, the central principle is the variation of the resonance signal over time. In these conditions, the spatial information may be maintained in a reference file, making it possible to compress subsequent images with no loss.

On the other hand, it is still necessary to evaluate the pertinence of the application of high compression rates when an assessment of structures of reduced dimensions relative to the size of the voxels is needed. Furthermore, the observation of the results from the application of the PCA technique in medical images may be considered a complexity measure.

In other words, images with dense texture patterns tend to produce different results with the use of the technique described. Nevertheless, this hypothesis was not tested in this project; it only points to the line of investigation, in which the results may certify and quantify this possibility. New secondary applications (based on the results here described) may encompass various conditions in the medical routine.

CHAPTER 7
CONCLUSION

7.1 Conclusion

Overall, in addition to its mathematical usages, linear algebra has broad usages and applications in most of engineering, medical, and biological field. As science and engineering disciplines grow so the use of mathematics grows as new mathematical problems are encountered and new mathematical skills are required. In this respect, linear algebra has been particularly responsive to computer science as linear algebra plays a significant role in many important computer science undertakings.

The broad utility of linear algebra to computer science reflects the deep connection that exists between the discrete nature of matrix mathematics and digital technology. In this thesis we have seen one important applications of the linear algebra which is called principal components analysis. This technique is used broadly in the medical field for compressing the medical images while keeping the good and needed features. However, this is not the only application of linear algebra in this field. Linear algebra has many other applications in this field. It provides many other concepts that are crucial to many areas of computer science, including graphics, image processing, cryptography, machine learning, computer vision, optimization, graph algorithms, quantum computation, computational biology, information retrieval and web search. Among these applications are face morphing, face detection, image transformations such as blurring and edge detection, image perspective removal, classification of tumors as malignant or benign, integer factorization, error-correcting codes, and secret-sharing.

REFERENCES

Kolman, B. (1996). Elementary Linear Algebra. New York, NY: Prentice Hall.

Sheldon, A. (2004). Linear Algebra Done Right. Germany: Springer.

Rajendra, B. (1996). Matrix Analysis, Graduate Texts in Mathematics. Germany: Springer.

Glazman, I. M. & Ljubic, Ju. I. (2006). Finite-Dimensional Linear Analysis. United Kingdom: Dover Publications.

Johnathan, S. G. (2007). The Linear Algebra a Beginning Graduate Student Ought to Know. Grmany: Springer.

Johnathan, S. G. (1995). Foundations of Linear Algebra. Netherlands: Kluwer.

Gerald, F., & Dianne H. (2004). Practical Linear Algebra: A Geometry Toolbox. United Kingdom: AK Peters.

Howard, A. (2005). Elementary Linear Algebra (Applications Version). USA: Wiley International.

David, C. (2005). Linear Algebra and Its Applications. USA: Addison Wesley.

Bernard, K., & David R. H. (2007). Elementary Linear Algebra with Applications. New York, NY: Prentice Hall.

Stephen H., Insel, A.J., & Lawrence, E. (2002). Linear Algebra. New York, NY: Prentice Hall.

Roger, A., & Charles, R. (1990). Matrix Analysis. United Kingdom: Cambridge University Press.

Thomas, S. (2006). Applied Linear Algebra and Matrix Analysis, Undergraduate Texts in Mathematics. Germany: Springer.

Larry, S. (1998). Linear Algebra, Undergraduate Texts in Mathematics. Germany: Springer.

Steven J. (2006), Linear Algebra with Applications. New York, NY: Pearson Prentice Hall.

David, P. (2010). Linear Algebra: A Modern Introduction. United Kingdom: Cengage – Brooks/Cole.

Henry, R. (2010). A Modern Introduction To Linear Algebra. New York, NY: CRC Press.

DeFranza, J., & Gagliardi, D. (2009). Introduction To Linear Algebra With Applications. New New York, NY: McGraw-Hill.

Sudipto, B., & Anindya, R. (2014). Linear Algebra and Matrix Analysis for Statistics, Texts in Statistical Science. USA: Chapman and Hall/CRC.

Gilbert, S. (2009). Introduction To Linear Algebra (4th ed.). United Kingdom: Wellesley-Cambridge Press.

Katta, G. (2014). Computational and Algorithmic Linear Algebra and n-Dimensional Geometry. USA: World Scientific Publishing.

Bretscher, O. (2004). Linear Algebra with Applications (3rd ed.). New York, NY: Prentice Hall.

Jolliffe, I.T. (1986). Principal Component Analysis. New York, NY: Springer-Verlag.

Mashal, N., Faust, M., & Hendler, T. (2005). The Role Of The Right Hemisphere In Processing Nonsalient Metaphorical Meanings: Application Of Principal Components Analysis To fMRI Data. *Neuropsychologia*, *43*(14), 284-300.

Gonzalez, R.C., & Woods, R.E. (1992). Digital Image Processing. Massachusetts: Addison-Wesley.

Boldrini, J.L., Costa, C.R., Figueirado, V.L., & Wetzler, H.G. (1984). Algebra Linear. São Paulo: Harbra.

Santo, R.D. (2012). Principal Component Analysis Applied To Digital Image Compression. *Einstein*, *10*(2), 135-140.