The Shape of Space

A Guided Tour of Vectors, Matrices, Tensors and Markov Chains

by **Stephen Fratini**

Table of Contents

List of Figures

List of Tables

Preface

I beseech you, in the bowels of Christ, think it possible that you may be mistaken.

Oliver Cromwell

The *Shape of Space: A Guided Tour of Vectors, Matrices, and Tensors* explores the rich structure and geometry that underpins modern linear algebra, with an emphasis on concepts that are both theoretically foundational and practically relevant. This book provides an accessible but rigorous treatment of a broad range of topics, including vector spaces, matrix decompositions, matrix calculus, tensors, non-negative matrices, Markov chains, and the least squares problem.

While some of these topics—such as vector spaces and linear transformations—are standard in a first course in linear algebra, others are less commonly included at the introductory level. In particular, this book ventures beyond the basics to cover matrix calculus, tensor analysis, and applications of non-negative matrices, including the Perron-Frobenius theorem and Markov chains. These extensions are motivated in part by their growing importance in fields such as machine learning, data science, and applied mathematics.

One of my primary motivations for authoring this book is to bridge the gap between the classical linear algebra curriculum and the mathematical tools needed to navigate contemporary applied work—especially in machine learning. Topics such as matrix calculus and tensors are now integral to modern computational models but are often introduced in a fragmented or overly technical manner. Here, I aim to present them in a unified and intuitive framework.

Throughout the book, I have aimed to balance clarity with precision, and conceptual insight with mathematical rigor. The goal is not just to present results but to foster an intuitive understanding of the structures that shape linear algebra and its many applications.

I hope this book will serve as a helpful resource for students, educators, and practitioners alike – especially those seeking a deeper understanding of the mathematical landscape underlying vectors, matrices, and tensors.

Legal

Stephen Fratini
Sole Proprietor of The Art of Managing Things
Eatontown, New Jersey (USA)
Email: sfratini@artofmanagingthings.com or sfratini@outlook.com
LinkedIn: www.linkedin.com/in/stephenfratini

Other books by the author:

- *The Art of Managing Things (2nd edition)*, self-published on Amazon, https://www.amazon.com/Art-Managing-Things-Stephen-Fratini-ebook/dp/B07N4H4YWH/, January 2019.
- *Mathematical Thinking: Exercises for the Mind (2nd Edition)*, self-published on Amazon, https://www.amazon.com/dp/B0CL34FRP1, October 2023.
- *Financial Mathematics (2nd Edition)*, self-published on Barnes and Noble, https://www.barnesandnoble.com/w/financial-mathematics-stephen-fratini/1145166826, March 2023.
- *Math in Art, and Art in Math*, self-published on Amazon, https://www.amazon.com/dp/B091D1F8MB, March 2021.
- *Algebra through Discovery and Experimentation*, self-published on Amazon, https://www.amazon.com/dp/B09B5L9WL5, July 2021.
- *The Struggle Against Chaos*, self-published on Amazon, https://www.amazon.com/dp/B09BLPQ86Q, July 2021.
- *Mathematical Vignettes: Number theory, stochastic processes, game theory, cryptography, linear programming and more*, self-published on Amazon, https://www.amazon.com/Mathematical-Vignettes-stochastic-cryptography-programming-ebook/dp/B0BBP1PBJQ/, August 2022.
- *Learning Math through Puzzles: Number properties, counting, sequences and series, algebra, functions, and mathematical reasoning*, self-published on Amazon, https://www.amazon.com/dp/B0BZFRZP5B, March 2023.
- *Mathematical Vignettes: Volume II: Topics from combinatorial design, magic squares, finite geometry, abstract algebra, error correcting codes, geometric packing problems and much more,* self-published on Amazon, *https://www.amazon.com/dp/B0CM1CLSK8, October 2023.*
- *Shape Up and Solve It!: Learn Geometry Through Puzzles*, self-published on Amazon, https://www.amazon.com/dp/B0CRS7DRWF, January 2024.
- *Mathematical Vignettes III (2nd edition): Introductions to non-Euclidean geometry, topology and complex analysis*, self-published on Amazon, https://www.amazon.com/Mathematical-Vignettes-III-Introductions-non-Euclidean/dp/B0F1N6522B, March 2025.

Electronic versions of my books are available (free of charge) at

https://github.com/sfratini33/art-of-managing-things-external/tree/master/free_books

Videos in support of this book:

- *Linear Algebra*, YouTube videos by the author that follows Section 2 of this book, https://www.youtube.com/playlist?list=PLSalWQAk7_UniISLwX82IFFqN1OsloWPK
- *Matrices - Some Less Common Properties,* YouTube videos by the author that follow Section 3 of this book, https://www.youtube.com/playlist?list=PLSalWQAk7_Umu0emwEcVtjDQs7O4ioZgr
- *Vector and Matrix Calculus*, YouTube videos by the author that follow Section 4 of this book, https://www.youtube.com/playlist?list=PLSalWQAk7_Ull8xLnrTOQaNe-KX6v0EDp
- *Tensors*, YouTube videos by the author that follow Sections 5 and 6 of this book, https://www.youtube.com/playlist?list=PLSalWQAk7_UltOlPJUX7nRHxIZ1RMYJXN
- *Non-negative matrices, Markov chains and Least Squares*, YouTube videos by the author that follow Section 7, 8 and 9 of this book, https://www.youtube.com/playlist?list=PLSalWQAk7_UlvuQ49JQUmV6GZGU7U1voD

Other video series by the author that may be of interest:

- *Puzzles to Exercise the Mind*, https://www.youtube.com/playlist?list=PLSalWQAk7_UmCBhNtgSpmgss0XAJq392j
- Non-Euclidean Geometry, https://www.youtube.com/playlist?list=PLSalWQAk7_UnXtByWwyjg9L-0JcA81HK8
- *Topology*, https://www.youtube.com/playlist?list=PLSalWQAk7_Uk62dBW9a6myMpmXrOllR90
- Complex Analysis, https://www.youtube.com/playlist?list=PLSalWQAk7_UkA9q_2Ptfz9mK9yzmngTRe

If you are working from a paper copy of my book, it may be easier to visit my YouTube channel rather than typing in the long URLs above.

www.youtube.com/@sfratini

1 Introduction

The greatest enemy of knowledge is not ignorance; it is the illusion of knowledge.

attributed to Stephen Hawking

1.1 Purpose

The purpose of this book is to bridge the gap between classical linear algebra and the mathematical tools increasingly used in modern applications, particularly in machine learning and data science. It's designed to give curious and motivated readers a clear, approachable path into both foundational ideas and more advanced topics without assuming a formal course or instructor. Whether you're brushing up or learning these topics for the first time, the goal is to make the material both understandable and useful.

1.2 Intended Audience

This book is intended for readers with a basic background in calculus, probability, and matrix operations who are looking to expand their understanding of linear algebra and its applications – particularly in areas related to machine learning. It is especially well-suited for self-study, offering a guided and structured approach for independent learners who want to build both intuition and technical fluency. Advanced undergraduates, graduate students, and professionals in fields such as computer science, engineering, and applied mathematics will find the material accessible and enriching, even without prior exposure to some of the more advanced topics covered.

1.3 Outline and Prerequisites

- Section 1 is this introduction.
- Section 2 provides an overview of some basic concepts of linear algebra. It is assumed the reader is familiar with how to solve a system of linear equations and the basic matrix operations (addition, multiplication and inversion).
- Section 3 covers some less common topics concerning matrices that are used in the following section on matrix calculus.
- Section 4 discusses matrix calculus (gradients, Jacobians, Hessians and other topics). It is assumed the reader has taken a college-level Calculus course.
- Section 5 provides a summary of the Einstein notation which we use in the following section on tensors.
- Section 6 is an introduction to tensors from two different (but equivalent) points of view, i.e., as multilinear maps and as multidimensional arrays.
- Section 7 introduces non-negative matrices with a focus on the Perron-Frobenius theorem.
- Section 8 is about Markov chains (an example of non-negative matrices). It is assumed the reader has taken a course in basic probability.
- Section 9 covers the least squares problem and its application to data fitting.

2 Linear Algebra

If you would be a real seeker after truth, it is necessary that at least once in your life you doubt, as far as possible, all things. – René Descartes

2.1 Vector Spaces

A **vector space** is a set whose elements (known as **vectors**) can be added together and multiplied ("scaled") by numbers called **scalars**. The operations of vector addition and scalar multiplication must satisfy certain requirements (known as vector axioms). Real vector spaces and complex vector spaces are kinds of vector spaces based on different kinds of scalars and vector elements, i.e., real numbers and complex numbers. More generally, scalars can be elements of a mathematical structure known as a field but the focus here is on the field of real numbers $\mathbb{R}$.

Formally, a vector space (over $\mathbb{R}$) consists of a set of vectors V along with the operations vector addition (represented by $+$) and scalar multiplication (represented by juxtaposition next to a vector). The two operations must meet the following axioms for all vectors $u, v, w \in V$ and all scalars $a, b \in \mathbb{R}$.

1. The set V is **closed under vector addition**, i.e., $u + v \in V$.
2. Vector addition is **commutative**, i.e., $u + v = v + u$.
3. Vector addition is **associative**, i.e., $u + (v + w) = (v + u) + w$.
4. There exists a **zero vector** $0 \in V$ such that $v + 0 = v$ for every $v \in V$.
5. Each $v \in V$ has an **additive inverse** $-v \in V$ such that $v + (-v) = 0$.
6. The set V is **closed under scalar multiplication**, $av \in V$.
7. Scalar multiplication **distributes** over scalar addition, i.e., $(a + b)v = av + bv$.
8. Scalar multiplication **distributes** over vector addition, i.e., $a(u + v) = au + av$.
9. Compatibility of scalar multiplication with vector multiplication, i.e., $(ab)v = a(bv)$.
10. Multiplication by the scalar 1 is the **identity** operation, i.e., $1v = v$.

[**Notation**: We will usually use the letters u, v, w for vectors, and we will use subscripting, if many vectors are needed, e.g., $v_1, v_2, \ldots, v_n$. For scalars, we will usually use the letters a, b, c and use subscripting if many scalars are needed. We will typically state at the beginning of a theorem, definition or discussion what variables are vectors, and which are scalars. When there is the possibility of confusion, we will use the notation $\vec{v}$ to represent the vector v.

In the above definitions, there are two types of addition, i.e., addition among vectors and common real number addition between scalars. Further, there are two types of multiplication, i.e., multiplication of a scalar times a vector and common real number multiplication between scalars.]

Example 1: Real Euclidean Plane, i.e., $\mathbb{R}^2$, over the real numbers $\mathbb{R}$

The real plane, i.e., the set $\mathbb{R}^2 = \{(x, y): x, y \in \mathbb{R}\}$, is a vector space where we define

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \text{ for } (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$$

$$a(x, y) = (ax, ay) \text{ for } (x, y) \in \mathbb{R}^2, a \in \mathbb{R}$$

Let's check that the axioms are fulfilled. [**Author's Remark**: Yes, this is a bit tedious, but please read through this carefully to reinforce your understanding of the axioms for a vector space.]

1. We have that $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \in \mathbb{R}^2$ since $x_1 + x_2, y_1 + y_2 \in \mathbb{R}$ since $\mathbb{R}$ is closed under addition, i.e., the sum of two real numbers is a real number.
2. We have that $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) = (x_2 + x_1, y_2 + y_1) = (x_2, y_2) + (x_1, y_1)$ which follows from the commutative in $\mathbb{R}$.
3. From the associative law in $\mathbb{R}$ (bold part of argument below), we have that
$$\begin{aligned}&(x_1, y_1) + \big((x_2, y_2) + (x_3, x_3)\big)\\ &= (x_1, y_1) + (x_2 + x_3, y_2 + y_3)\\ &= \big(x_1 + (x_2 + x_3), y_1 + (y_2 + y_3)\big) = \big((x_1 + x_2) + x_3, (y_1 + y_2) + y_3\big)\\ &= (x_1 + x_2, y_1 + y_2) + (x_3, y_3)\\ &= \big((x_1, y_1) + (x_2, y_2)\big) + (x_3, y_3)\end{aligned}$$
4. The zero vector is $(0,0)$ since for any vector $(x, y) \in \mathbb{R}^2$, we have $(x, y) + (0,0) = (x + 0, y + 0) = (x, y)$.
5. The additive inverse for a vector $(x, y) \in \mathbb{R}^2$ is $(-x, -y)$, since $(x, y) + (-x, -y) = (x - x, y - y) = (0,0)$.
6. Since $\mathbb{R}$ is closed under multiplication, we have $a(x, y) = (ax, ay) \in \mathbb{R}^2$.
7. From our definition for addition and scalar multiplication in $\mathbb{R}^2$, and the distribute law for $\mathbb{R}$, we have
$$\begin{aligned}&(a + b)(x, y) = \big((a + b)x, (a + b)y\big)\\ &= (ax + bx, ay + by) = (ax, ay) + (bx, by)\\ &= a(x, y) + b(x, y)\end{aligned}$$
8. From our definitions for addition and scalar multiplication in $\mathbb{R}^2$, and the distribute law for $\mathbb{R}$ we have
$$\begin{aligned}&a\big((x_1, y_1) + (x_2, y_2)\big) = a(x_1 + x_2, y_1 + y_2)\\ &= \big(a(x_1 + x_2), a(y_1 + y_2)\big) = (ax_1 + ax_2, ay_1 + ay_2)\\ &= (ax_1, ay_1) + (ax_2, ay_2) = a(x_1, y_1) + a(x_2, y_2)\end{aligned}$$
9. From the associative law in $\mathbb{R}$, we have that $(ab)(x, y) = \big((ab)x, (ab)y\big) = \big(a(bx), a(by)\big) = a(bx, by) = a\big(b(x, y)\big)$.
10. Finally, $1(x, y) = (1x, 1y) = (x, y)$ since $1 \in \mathbb{R}$ times a number is the number itself.

Example 2: n-dimensional Euclidean space $\mathbb{R}^n$ over $\mathbb{R}$

We can extend the previous example to n dimensions. First, we define the set

$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_i \in \mathbb{R}, i = 1,2, \dots, n\}$.

Addition is done pairwise, i.e.,

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

Scalar multiplication is defined as follows:

$$a(x_1, x_2, \dots, x_n) = (ax_1, ax_2, \dots, ax_n), a \in \mathbb{R}$$

Verification of the vector space axioms follows a similar approach to what we did for $\mathbb{R}^2$.

$\mathbb{R}^n$ is formed as a **Cartesian product** of n copies of $\mathbb{R}$. We write this as $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R} = \{(x_1, x_2, \dots, x_n): x_i \in \mathbb{R}, i = 1,2, \dots, n\}$. In general, we can form the Cartesian product of any collection of sets.

There are several variations to this example, i.e.,

- $\mathbb{R}^\infty = \{(x_1, x_2, \dots): x_i \in \mathbb{R}, i = 1,2, \dots\}$, with pairwise addition and scalar multiplication similar to what we did for $\mathbb{R}^n$, forms an infinite dimensional vector space. This is sometimes denoted as $\mathbb{R}^\mathbb{N}$ where $\mathbb{N}$ is the set of natural numbers $\{1,2,3, \dots\}$.
- $\mathbb{R}^n$ with its standard scalar multiplication rule over the set of complex numbers $\mathbb{C}$ is not a vector space, since the axiom concerning closure under scalar multiplication does not hold, e.g., $i(1,0) = (i,0)$ is not in $\mathbb{R}^2$.
- n-dimensional complex space $\mathbb{C}^n$ over $\mathbb{R}$ is a valid vector space.
- n-dimensional complex space $\mathbb{C}^n$ over $\mathbb{C}$ is a valid vector space.

Example 3: Consider the set L all scalar multiples of a given constant vector $w = (c_1, c_2, \dots, c_n) \in R^n, n \geq 2$. L is a vector space (or more precisely, a subspace) within the vector R^n.

Let's check that the vector space axioms hold for L.

1. Take $u, v \in L$. This means that $u = aw$ and $v = bw$ for some $a, b \in \mathbb{R}$. Adding the two vectors, we have $u + v = aw + bw = (a + b)w \in L$. Note that $aw + bw = (a + b)w$ follows from the distributive law for $\mathbb{R}^n$ which L inherits from $\mathbb{R}^n$ since it is a subset of $\mathbb{R}^n$.
2. L inherits this property from $\mathbb{R}^n$.
3. L inherits this property from $\mathbb{R}^n$.
4. $0w = \vec{0} \in L$. Note that we have the scalar 0 and the vector $\vec{0}$ in the previous expression.
5. If $u = aw \in L$ then $(-aw) \in L$ is the additive inverse of u since $aw + (-aw) = (a - a)w = 0w = \vec{0}$.
6. If $u = aw \in L$, then $b(aw) = (ba)w \in L$ and so, we have closure by multiplication of a scalar.
7. L inherits this property from $\mathbb{R}^n$.
8. L inherits this property from $\mathbb{R}^n$.
9. L inherits this property from $\mathbb{R}^n$.
10. L inherits this property from $\mathbb{R}^n$.

L is a straight line in $\mathbb{R}^n$ that passes through the origin $\vec{0} = (0,0, \dots,0)$. Lines in $\mathbb{R}^n$ that don't pass through the origin are not vector spaces.

Example 4: The set of all polynomial of order $n-1$ with real coefficients, i.e., $\mathcal{P}_n = \{a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} : a_i \in \mathbb{R}, i = 1,2,\ldots,n-1\}$, is a vector space of size n under the usual definition of polynomial addition and multiplication by a scalar. In fact, $\mathcal{P}_n$ is the same as $\mathbb{R}^n$ with regard to vector space properties.

Example 5: The set of all real-valued matrices of size $m \times n$ is a vector space. Addition is defined pairwise, i.e., add element ij in one matrix to element ij in another matrix to get entry ij in the sum of the two matrices. Scalar multiplication is applied to each entry in a matrix. This vector space is the same as $\mathbb{R}^{mn}$ in terms of vector space properties.

In the following, we show an example of adding two 3×3 matrices, and multiplication by a scalar.

$$\begin{pmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 2 & 4 & 4 \\ -1 & 1 & 7 \\ 9 & 5 & -3 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 7 \\ 5 & 6 & 11 \\ 16 & 13 & 6 \end{pmatrix}$$

$$3\begin{pmatrix} 1 & 7 & 2 \\ 5 & 1 & 6 \\ 2 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 21 & 6 \\ 15 & 3 & 18 \\ 6 & 12 & 3 \end{pmatrix}$$

Example 6: The set F of all function from the natural numbers $\mathbb{N}$ to the real numbers $\mathbb{R}$ is a vector space under the usual operations for function addition and multiplication by a scalar. Each mapping can be represented as a vector in $\mathbb{R}^\infty$. It is left as an exercise for the reader to check that the vector space axioms hold true for F, e.g.,$f(n) = 0, n = 1,2,\ldots$ is the zero vector.

Further, consider $f(n) = n^2 + \pi$. The mappings for this function are represented in the following table:

n	1	2	3	4	5	…
$f(n)$	π	$4+\pi$	$9+\pi$	$16+\pi$	$25+\pi$	…

This can, in turn, be represented by the vector $v_f = (\pi, 4+\pi, 9+\pi, 16+\pi, 25+\pi, \ldots)$. The set of all scalar multiplies of v_f is a subspace of the vector space F (basically, a straight line in infinite dimensional Euclidean space $\mathbb{R}^\infty$).

Example 7: The set of solutions of a homogeneous linear system $Ax = 0$ in n variables is a vector space under the operations inherited from $\mathbb{R}^n$. In notation, we are talking about the following set:

$$S = \{x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n : Ax = 0\}$$

[The above expression $(x_1, x_2, \ldots, x_n)^T$ effectively maps the row vector $(x_1, x_2, \ldots, x_n)$ to the column vector x.]

To show additive closure, assume take $u, v \in S$ which implies $Au = 0$ and $Av = 0$. We have that

$$A(u+v) = Au + Av = 0 + 0 = 0$$

Thus, $u + v \in S$.

Showing closure under scalar multiplication is even easier. If $u \in S, \alpha \in \mathbb{R}$ then $A(\alpha u) = \alpha A u = \alpha 0 = 0$ and thus, $\alpha u \in S$.

The vector $\vec{0} = (0,0,\ldots,0)$ is always a solution and thus a member of the set.

Checking the other axioms is straightforward and left as an exercise.

[Later on in this book, we will see that if columns of A are something known as linearly independent, then there is only one solution to $Ax = 0$, i.e., $x = 0$. In that case, the vector space is not very interesting.]

The following theorem provides some very basic properties for vector spaces.

Theorem 1. Given any vector space V, for any $v \in V$ and scalar $a \in \mathbb{R}$, the following results hold true

(1) $\mathbf{0}v = \vec{\mathbf{0}}$

(2) $(-\mathbf{1}v) + v = \vec{\mathbf{0}}$***, i.e.,*** $-\mathbf{1}v = -v$

(3) $a\vec{\mathbf{0}} = \vec{\mathbf{0}}$***.***

Proof: Note that we are using the vector notation here to distinguish between the scalar 0 and the vector $\vec{0}$.

For property (1), we first note that $v = (1 + 0)v = v + 0v$. Let u be the additive inverse of v and add v to both sides of the equation to get

$$u + v = u + v + 0v$$
$$\vec{0} = \vec{0} + 0v$$
$$\vec{0} = 0v$$

For property (2), we have $(-1v) + v = (-1v) + (1v) = (-1 + 1)v = 0v = \vec{0}$.

For property (3), we have $a\vec{0} = a(0\vec{0}) = (a0)\vec{0} = 0\vec{0} = \vec{0}$. ■

2.2 Subspaces and Spanning Sets

A **subspace** of a vector space is a subset of the vectors in a vector space that also form a vector space. The empty set is defined to be a subspace, and is known as the trivial subspace. We already saw an example of a subspace in Example 3 from the previous section, i.e., the line formed by all scalar multiples of a vector is a subspace of $\mathbb{R}^n$.

Example 1: For $m < n$, $\mathbb{R}^m$ is a subspace of $\mathbb{R}^n$. This follows since we know $\mathbb{R}^m$ is a vector space and clearly, it is a subset of $\mathbb{R}^n$.

Example 2: Consider the set $P = \{(x, y, z) \in \mathbb{R}^3 : x - \frac{1}{2}y + z = 0\}$. By definition, P is a subset of $\mathbb{R}^3$.

Solving for y in terms of x and z, we have $y = 2x + 2z$. So, the elements of P can be represented as follows:

$$(x, y, z) = (x, 2x + 2z, z) = x(1,2,0) + z(0,2,1)$$

So, P is the set of all linear combinations of the vectors $(1,2,0)$ and $(0,2,1)$. The set P is a plane through the origin in $\mathbb{R}^3$.

Most of the vector space properties are inherited from $\mathbb{R}^3$, but we should verify the existence of inverses and the zero element, and the closure axioms.

Clearly, $\vec{0} \in P$.

It is also clear that additive inverses exist for each $(x, y, z) \in P$, i.e., if $(x, y, z) \in P$ then so is $(-x, -y, -z) = -x(1,2,0) - y(0,2,1)$.

In terms of additive closure, the sum of two vectors that are linear combinations of $(1,2,0)$ and $(0,2,1)$ can also be represented as a linear combination of $(1,2,0)$ and $(0,2,1)$.

Closure for scalar multiplication also holds, i.e., $a(x, y, z) = ax(1,2,0) + az(0,2,1) \in P$ if $(x, y, z) \in P$.

Example 3. Let H be the set $\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 + x_2 + x_3 - x_4 = 0\}$. By definition, H is a subset of $\mathbb{R}^4$.

Solving for x_4 in terms of the other three variables, we have that $x_4 = x_1 + x_2 + x_3$. So, the elements of H can be written in the form

$$(x_1, x_2, x_3, x_4) = (x_1, x_2, x_3, x_1 + x_2 + x_3) = x_1(1,0,0,1) + x_2(0,1,0,1) + x_3(0,0,1,1)$$

Thus, H consists of all linear combinations of the vectors $(1,0,0,1)$, $(0,1,0,1)$ and $(0,0,1,1)$. H is a **hyperplane** in $\mathbb{R}^4$. We leave it as an exercise for the reader to verify that H is a subspace of $\mathbb{R}^4$.

[From Wikipedia: "In geometry, a hyperplane is a generalization of a two-dimensional plane in three-dimensional space to mathematical spaces of arbitrary dimension. Like a plane in space, a hyperplane is a flat hypersurface, **a subspace whose dimension is one less than that of the ambient space**. Two lower-dimensional examples of hyperplanes are one-dimensional lines in a plane and zero-dimensional points on a line."]

Example 4. For any vector space V, the set $\{\vec{0}\}$ is always a subspace.

Exercise 1: Verify that the set $\{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 : x_1 + x_2 - 2x_3 + x_4 + x_5 = 0\}$ is a subspace of $\mathbb{R}^5$.

Exercise 2: Why is the set $\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 + x_2 + x_3 - x_4 = 1\}$ **not** a subspace of $\mathbb{R}^4$.

The following theorem simplifies the task of determining whether a subset of a vector space is a subspace.

Theorem 2. For a nonempty subset S of a vector space V, under the inherited operations the following are equivalent statements.

- ***S is a subspace of V.***
- ***S is closed under linear combinations of pairs of vectors, i.e., for any vectors $u, v \in S$ and scalars $a, b \in \mathbb{R}$, $au + bv \in S$.***

Proof: If S is a subspace of V, all linear combinations of vectors in S are also in S. This follows from the two closure axioms, i.e., Axioms #1 and #6.

Going in the other direction, assume that S is closed under linear combinations. That immediately gives us Axioms #1 and #6.

Concerning Axiom #4, we have that $0u + 0u = \vec{0} + \vec{0} = \vec{0}$.

If $u \in S$, then $-1u = -u \in S$ (making use of $\text{Theorem } 1$). So, Axiom #5 holds true.

Axioms 2,3,7,8,9 and 10 are all inherited by S from V. ■

Example 5. Consider the following set

$$M = \left\{ \begin{pmatrix} x & 0 \\ y & z \end{pmatrix} : x, y, z \in \mathbb{R}, x + y - z = 0 \right\}$$

We can write $z = x + y$ and then decompose the elements of M as the linear combination of two matrices, i.e.,

$$\begin{pmatrix} x & 0 \\ y & z \end{pmatrix} = \begin{pmatrix} x & 0 \\ y & x+y \end{pmatrix} = x \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

By Theorem 2, we only need to show that M is closed under linear combinations. Taking two elements of M (using the above format), we have

$$a \begin{pmatrix} x_1 & 0 \\ y_1 & z_1 \end{pmatrix} + b \begin{pmatrix} x_2 & 0 \\ y_2 & z_2 \end{pmatrix}$$

$$= a \left[x_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y_1 \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \right] + b \left[x_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y_2 \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \right]$$

$$= (ax_1 + bx_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (ay_1 + by_2) \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \in M$$

$$\ldots$$

The **span** of a nonempty subset $S = \{s_1, s_2, \ldots, s_n\}$ of a vector space is the set of all linear combinations of vectors from S, i.e.,

$$\text{span}(S) = \{a_1 s_1 + a_2 s_2 + \cdots + a_n s_n : s_i \in S, a_i \in \mathbb{R}, i = 1,2, \ldots, n\}$$

In the case of an empty set, we define its span to be $\{\vec{0}\}$.

In Example 3 of this section, we showed that the span of $\{(1,0,0,1), (0,1,0,1), (0,0,1,1)\}$ is the set H.

Theorem 3. The span of any subset of a vector space is a subspace of the vector space, and conversely, i.e., every subspace is the span of some set.

Proof: If the subset S is empty then by definition it spans the trivial subspace.

If S is non-empty then by Theorem 2, we need only check that $\text{span}(S)$ is closed under linear combinations of pairs of elements. Let $S = \{s_1, s_2, \ldots, s_n\}$. Select two vectors from $\text{span}(S)$, i.e.,

$$v = a_1 s_1 + a_2 s_2 + \cdots + a_n s_n, \qquad u = b_1 s_1 + b_2 s_2 + \cdots + b_n s_n$$

Take any linear combination of v and u as shown below. Note that some of the a_i or b_i can equal zero.

$$xv + yu = x(a_1 s_1 + a_2 s_2 + \cdots + a_n s_n) + y(b_1 s_1 + b_2 s_2 + \cdots + b_n s_n), \qquad x, y \in \mathbb{R}$$

$$(xa_1 + yb_1)s_1 + (xa_2 + yb_2)s_2 + \cdots + (xa_n + yb_n)s_n \in \text{span}(S)$$

Going in the other direction, the vectors of a subspace span the subspace. This is true because of the two closure axioms. ■

Theorem 4. The span of a subset S of a vector space V is the smallest subspace containing all of the members of S.

Proof: From Theorem 3, we know that $span(S)$ is a subspace of V.

Next, we show that $span(S)$ is the smallest such subspace.

Let W be any subspace of V that contains S. We need to show that $span(S) \subseteq W$.

Since W is a subspace and $S \subseteq W$, W must contain all linear combinations of vectors in S (because subspaces are closed under linear combinations). Thus, $span(S) \subseteq W$.

This shows that $span(S)$ is contained in every subspace that contains S, making it the smallest such subspace. ■

Example 6. In any vector space V and vector $v \in V$, the set $L = \{av: a \in \mathbb{R}\}$ is a subspace of V. This follows from Theorem 3 since L is (by definition) the span of v. We can think of L as sort of a generalized line.

Example 7. The span of vectors $(1,0)$ and $(0,1)$ is all of $\mathbb{R}^2$. This is easy to see since we can write any vector $v = (x, y) \in \mathbb{R}^2$ as follows:

$$(x, y) = x(1,0) + y(0,1)$$

Some obvious questions:

- Is two the smallest number of vectors that span all of $\mathbb{R}^2$?
- Assuming the answer to the first question is "yes", are their other pairs of vectors that span $\mathbb{R}^2$? This question has to do with something known as linear independence, which we will discuss in the next section.

2.3 Linear Independence

In a given vector space, a set of vectors is said to be **linearly independent** if none of its elements is a linear combination of the others from the set; otherwise, the set is said to be linearly dependent.

For example, the vectors $(1,0)$ and $(0,1)$ are linearly independent in $\mathbb{R}^2$ since there is no way of writing one of the vectors as a linear combination of the other. On the other hand, the vectors $(1,0,0)$, $(0,1,0)$ and $(2,2,0)$ are linearly dependent in $\mathbb{R}^3$ since we can write

$$(2,2,0) = 2(1,0,0) + 2(0,1,0)$$

Theorem 5. Given a subset S of a vector space V and some vector $v \in V$, $span(S \cup \{v\}) = span(S)$ if and only if $v \in S$.

Proof: If $v \in S$, then $S \cup \{v\} = S$ and thus, $span(S \cup \{v\}) = span(S)$.

Assume $span(S \cup \{v\}) = span(S)$ and let $S = \{s_1, s_2, \dots, s_n\}$. Clearly, $v \in span(S \cup \{v\}) = span(S)$ which implies the v can be written as a linear combination of the vectors in S, i.e.,

$$v = a_1 s_1 + a_2 s_2 + \cdots + a_n s_n, \qquad a_i \in \mathbb{R}, i = 1,2, \dots, n$$

If $v \notin S$, then v is a linear combination of elements of S but is not itself in S. However, if v were not in S, then adding v to S would imply that $span(S \cup \{v\})$ has a larger span than $span(S)$, contradicting our assumption that $span(S \cup \{v\}) = span(S)$. Thus, $v \in S$. ■

Theorem 5 implies the following result.

Theorem 6. Given $v \in S \subseteq V$, where V is a vector space. $span(S) = span(S - \{v\})$ if and only if v is linearly dependent on the other vectors of S.

In other words, if $v \in S$ can be written as a linear combination of the other vectors in S then removing v from S does not change (decrease) the span of S.

Theorem 7. A subset $S = \{s_1, s_2, \dots, s_n\}$ of a vector space V is linearly independent if and only if $a_1 s_1 + a_2 s_2 + \cdots + a_n s_n = 0$ implies that $a_1 = a_2 = \cdots = a_n = 0$.

Proof: If S is not linearly independent, then some s_i can be written as a linear combination of the other vectors in S, i.e.,

$$s_i = a_1 s_1 + \cdots + a_{i-1} s_{i-1} + a_{i+1} s_{i+1} + \cdots + a_n s_n$$

$$\Rightarrow\ a_1 s_1 + \cdots + a_{i-1} s_{i-1} - 1 s_i + a_{i+1} s_{i+1} + \cdots + a_n s_n = 0$$

Thus, we have a linear combination of the vectors in S equal to 0 with at least one coefficient not equal to zero (i.e., $a_i = -1$).

On the other hand, if S is linearly independent then (by definition) no vector $s_i \in S$ is a linear combination of other vectors in S. Thus, there can be no linear relationship among the vectors in S where some of the coefficients have a nonzero value. ■

Example 1. Determine whether the vectors $(1,2,3), (4,7,2), (9,3,1)$ are linearly independent in $\mathbb{R}^3$.

To solve the problem, we need to determine whether the following system of equations has any non-zero solutions:

$$x_1(1,2,3) + x_2(4,7,2) + x_3(9,3,1) = \vec{0}$$

The above can be written as a linear system of three equations in three unknowns, i.e.,

$$x_1 + 4x_2 + 9x_3 = 0$$

$$2x_1 + 7x_2 + 3x_3 = 0$$

$$3x_1 + 2x_2 + x_3 = 0$$

We can write this as the matrix equation $A\vec{x} = \vec{0}$ where

$$A = \begin{pmatrix} 1 & 4 & 9 \\ 2 & 7 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

If the inverse of matrix A exist, then we have the unique solution $\vec{x} = \vec{0}$ to our matrix equation. This would imply that the given vectors are linearly independent. In fact, the inverse of A exists, i.e.,

$$A^{-1} = \frac{1}{124}\begin{pmatrix} -1 & -7 & 17 \\ -14 & 26 & -10 \\ 51 & -15 & 1 \end{pmatrix}$$

Computation of the above inverse matrix was done using Wolfram Alpha with the command

inv {{1,4,9},{2,7,3},{3,2,1}}

Another option is to use the Julia package from MIT, see https://docs.julialang.org/en/v1/

```
julia> A=[1 4 9;2 7 3;3 2 1]

3×3 Matrix{Int64}:

 1  4  9
 2  7  3
 3  2  1

julia> A^-1

3×3 Matrix{Float64}:

 -0.00806452  -0.112903    0.41129
 -0.0564516    0.209677   -0.120968
  0.137097    -0.0806452   0.00806452
```

Example 2. Determine whether the vectors (1,2,3), (7,9,2),(8,11,5) are linearly independent in $\mathbb{R}^3$. To solve the problem, we need to determine whether the following system of equations has any non-zero solutions:

$$x_1(1,2,3) + x_2(7,9,2) + x_3(8,11,5) = \vec{0}$$

The above can be written as a linear system of three equations in three unknowns, i.e.,

$$x_1 + 7x_2 + 8x_3 = 0$$

$$2x_1 + 9x_2 + 11x_3 = 0$$

$$3x_1 + 2x_2 + 5x_3 = 0$$

We could write the above in the form $A\vec{x} = \vec{0}$ but in this case, A^{-1} does not exist (meaning there is not a single solution to the system). So, we will form the augment matrix (first matrix on the left below) and transpose it into reduced row echelon form.

$$\left(\begin{array}{ccc|c} 1 & 7 & 8 & 0 \\ 2 & 9 & 11 & 0 \\ 3 & 2 & 5 & 0 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 7 & 8 & 0 \\ 0 & -5 & -5 & 0 \\ 0 & -19 & -19 & 0 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 7 & 8 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Thus, we have that $x_1 = -x_3$ and $x_2 = -x_3$. The system of equations has an infinite number of solutions of the form

$$(x_1, x_2 x_3) = (-x_3, -x_3, x_3) = x_3(-1, -1, 1)$$

In conclusion, the three given vectors are linearly dependent.

Another approach is to determine the span of (1,2,3), (7,9,2),(8,11,5). If it is not all of $\mathbb{R}^3$ then the vectors are dependent. Note that the span of the three vectors is not affected by row operations. So, in the last step below (on the right), we can remove $(0,0,0)$ and not decrease the span. Thus, by Theorem 6, the initial set of vectors is dependent.

$$\begin{pmatrix} 1 & 2 & 3 \\ 7 & 9 & 2 \\ 8 & 11 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -5 & -19 \\ 0 & -5 & -19 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & \frac{19}{5} \\ 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -\frac{23}{5} \\ 0 & 1 & \frac{19}{5} \\ 0 & 0 & 0 \end{pmatrix}$$

So, the give three vectors span a plane in $\mathbb{R}^3$ and are thus dependent. The said plane is spanned by the vectors $(5,0,-23)$ and $(0,5,19)$.

Example 3. In any vector space, any subset containing the zero vector is linearly dependent. For example, take the subset $\{(0,0,0),(1,0,0)\}$ in the vector space $\mathbb{R}^3$. We can write $(0,0,0) = 0(1,0,0)$ and thus, the two vectors are linearly dependent.

Example 4. The empty subset of a vector space is linearly independent, since there is no nontrivial linear relationship among its members of which there are none. So, the linear independence property is fulfilled vacuously.

As a convention, we define the sum of zero-many vectors to be the zero vector. This seems odd but it will be needed in some of the following discussions.

The following theorem states that a spanning set is minimal if and only if it is linearly independent.

Theorem 8. A set S is linearly independent if and only if for any vector $v \in S$, its removal shrinks the span, i.e., $span(S - \{v\}) \subset span(S)$.

Note that we used $\subset$ versus $\subseteq$ to indicate a proper subset which excludes the possibility of equal.

Proof: If S is linearly independent and we remove some vector v, then there is no way to represent v with the remaining vectors in S. Thus, $span(S - \{v\}) \subset span(S)$.

If S is linearly dependent, then there is some vector $v \in S$ such that v is a linear combination of the other vectors in S. So, if we remove v, we have that $span(S - \{v\}) = span(S)$. ■

The following theorem concerns adding a vector to a set.

Theorem 9. Given a linearly independent set S in a vector space V and a vector $v \notin S$. The set $S \cup \{v\}$ is linearly independent if and only if $v \notin span(S)$.

The theorem is equivalent to the following: "The set $S \cup \{v\}$ is linearly dependent if and only if $v \in \text{span}(S)$." We will prove this form of the theorem.

Proof: If $v \in \text{span}(S)$, then clearly, $S \cup \{v\}$ is linearly dependent.

Going the other way, assume $S \cup \{v\}$ is linearly dependent. This implies there is a linear relationship among the elements of $S \cup \{v\}$ with one or more of the coefficients being nonzero, i.e.,

$$a_1 s_1 + a_2 s_2 + \cdots + a_n s_3 + av = \vec{0}, \qquad s_i \in S, a_i \in \mathbb{R}$$

If $a = 0$, then at least one of the $a_i = 0$ but that implies that S is linearly dependent (a contradiction). Thus, $a \neq 0$. So, we can rewrite the above equation as

$$v = -\frac{a_1}{a} s_1 - \frac{a_2}{a} s_2 - \cdots - \frac{a_n}{a} s_n$$

This implies that $v \in \text{span}(S)$. ■

The following theorem, which we state without proof, summarizes some of the previous discussion.

Theorem 10. A subset of a linearly independent set is also linearly independent. A superset of a linearly dependent set is also linearly dependent.

2.4 Basis

A **basis** for a vector space V is a set of vectors from V that is linearly independent and that spans the space.

Some definitions of "basis" consider the order of the elements and define a basis as a "sequence of vectors" rather than "set of vectors". We will revisit this point when we get to the representation of a vector in terms of a basis.

For example, $\{(1,0,0),(0,1,0),(0,0,1)\}$ is a basis for $\mathbb{R}^3$.

The following is known as the **standard basis** for $\mathbb{R}^n$:

$$\{e_1 = (1,0,\dots,0), e_2 = (0,1,0,\dots,0), \dots, e_n = (0,0,\dots,0,1)\}$$

To be clear, e_i has a 1 in position i and a zero in all the other positions. In terms of notation, we will refer to this basis as $\mathcal{E}_n$.

For any $v = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, we can write v in terms of the standard basis as follows:

$$v = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

Theorem 11. A subset S of a vector space V is a basis if and only if each vector in V can be expressed <u>uniquely</u> as a linear combination of vectors in S.

Proof: Assume $S = \{s_1, s_2, \dots, s_n\}$ is a basis. By way of contradiction, assume there exist some $v \in V$ that can be represented in two different ways as a linear combination of the vectors in S, i.e.,

$$v = a_1 s_1 + a_2 s_2 + \cdots + a_n s_n = b_1 s_1 + b_2 s_2 + \cdots + b_n s_n$$

Subtracting the two representations, we get

$$(a_1 - b_1)s_1 + (a_2 - b_2)s_2 + \cdots + (a_n - b_n)s_n = \vec{0}$$

Since the vectors of a basis are (by definition) linearly independent, it must be that $a_i = b_i, i = 1,2,\dots,n$ and so, the two representations are the same.

Going the other way, assume every vector in V can be expressed uniquely as a linear combination of vectors in S. This implies that S spans V. So, we only need to prove the vectors in S are linearly independent. Suppose there is a nontrivial linear dependence among the vectors in S, where at least one coefficient in the following equation does not equal 0 (say a_i):

$$a_1 s_1 + a_2 s_2 + \cdots + a_n s_n = \vec{0}, a_i \neq 0$$

This gives us two ways to express the zero vector $\vec{0}$ as a linear combination of basis vectors, i.e., one equation with all coefficients equal to zero, and another using the above equation where $a_i \neq 0$. This contradicts the assumption that each vector (including the zero vector) has a unique representation. ■

In a vector space V with basis $S = \{s_1, s_2, \dots, s_n\}$ the representation of $v \in V$ with respect to S is the vector of the coefficients used to express v as a linear combination of the vectors in S. We write this as

$$Rep_S(v) = (a_1, a_2, \dots, a_n)_S$$

where $v = a_1 s_1 + a_2 s_2 + \cdots + a_n s_n$.

Alternatively, we can use a column vector, i.e.,

$$Rep_S(v) = \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix}$$

If we change the order of the vectors in S, then the representation of a given vector changes. This is why some authors define a basis as a sequence (i.e., an ordered list).

Example 1. The set $S = \{(2,0,0), (0,3,0), (0,0,4)\}$ is a basis for $\mathbb{R}^3$. The representation vector for $(2,2,2)$ under this basis is $\left(1, \frac{2}{3}, \frac{1}{2}\right)_S$ since $(2,2,2) = (2,0,0) + \frac{2}{3}(0,3,0) + \frac{1}{2}(0,0,4)$.

Example 2. Consider the vector space of all real 2×2 matrices. The following set is a basis:

$$M = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

The representation for $\begin{pmatrix} 2 & 3 \\ 7 & 1 \end{pmatrix}$ in the basis M is $(2,3,7,1)_M$.

Theorem 12. Let S be a basis for a vector V, where S has n elements. The equation $\boldsymbol{a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = \vec{0}}$ ***for*** $\boldsymbol{v_j \in V, a_j \in \mathbb{R}, j = 1, 2, \dots, k}$ ***holds true if and only if*** $\boldsymbol{a_1 Rep_S(v_1) + a_2 Rep_S(v_2) + \cdots + a_k Rep_S(a_k) = \vec{0}_S}$ ***holds true.***

Proof: Let $S = \{s_1, s_2, \dots, s_n\}$.

Form the $n \times k$ matrix $C = [c_{ij}]$ such that column i of C is $Rep_S(v_i)$, i.e.,

$$v_j = c_{1j} s_1 + c_{2j} s_2 + \cdots + c_{nj} s_n, \qquad j = 1,2, \dots, k$$

We have that

$$\vec{0} = a_1 v_1 + \cdots + a_k v_k$$

$$\Leftrightarrow a_1(c_{11} s_1 + c_{21} s_2 + \cdots + c_{n1} s_n) + \cdots + a_k(c_{1k} s_1 + c_{2k} s_2 + \cdots + c_{nk} s_n) = \vec{0}$$

$$\Leftrightarrow (a_1 c_{11} + \cdots + a_k c_{1k}) s_1 + \cdots + (a_1 c_{n1} + \cdots + a_k c_{nk}) s_n = \vec{0}$$

Since the S is a basis, its vectors are linearly independent which implies that the coefficients in the equation above must be equal to 0, i.e.,

$$a_1c_{11} + \cdots + a_kc_{1k} = 0$$
$$a_1c_{21} + \cdots + a_kc_{2k} = 0$$
$$\ldots$$
$$a_1c_{n1} + \cdots + a_kc_{nk} = 0$$

This is equivalent to the equation

$$a_1Rep_S(v_1) + a_2Rep_S(v_2) + \cdots + a_kRep_S(a_k) = \vec{0}$$

The above flow follows in both the forward and reverse directions with equivalence at each step. ■

Example 3. Take $(1,1,1), (2,2,0), (0,0,2) \in \mathbb{R}^3$. We have that

$$-2(1,1,1) + (2,2,0) + (0,0,2) = \vec{0}$$

Let's use the basis $S = \{(2,0,0), (0,3,0), (0,0,4)\}$ and apply Theorem 12. First, we represent our three vectors in terms of basis S, i.e.,

$$Res_S(1,1,1) = \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{4}\right)_S, \qquad Res_S(2,2,0) = \left(1, \frac{2}{3}, 0\right)_S, \qquad Res_S(0,0,2) = \left(0,0, \frac{1}{2}\right)_S$$

Next, we verify the result of the theorem holds true here, i.e.,

$$-2Res_S(1,1,1) + Res_S(2,2,0) + Res_S(0,0,2) = -2\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{4}\right)_S + \left(1, \frac{2}{3}, 0\right) + \left(0,0, \frac{1}{2}\right)_S$$

$$= \left(-1, -\frac{2}{3}, -\frac{1}{2}\right)_S + \left(1, \frac{2}{3}, 0\right) + \left(0,0, \frac{1}{2}\right)_S = (0,0,0)_S$$

2.5 Dimension

The **dimension of a vector space** is the number of vectors in a basis for the vector space. However, for this definition to have meaning, we first need to prove that every basis for a given finite dimensional vector space has the same number of vectors.

First, we define a finite-dimensional vector space as one that has **a** basis with only finitely many vectors. Our focus in this section will be on such vector spaces.

A **trivial space** is a vector space with only the zero vector. The trivial vector space has dimension 0, since it has no basis (or equivalently, the empty set ϕ is its basis, since the span of the empty is $\{\vec{0}\}$).

To prove that every basis for a given finite dimensional vector space has the same number of vectors, we first need to prove the following intermediate result.

Theorem 13. Assume that $S = \{s_1, s_2, \dots, s_n\}$ is a basis for a vector space V, and that the vector $v \in V$ can be represented as $v = a_1 s_1 + a_2 s_2 + \cdots + a_n s_n$ where $a_i \neq 0$. Then exchanging s_i for v yields another basis for V.

Proof: Let S' be the set S with s_i replaced by v.

We first show that the vectors in S' are linearly independent. Taken any linear combination of the vectors in S' that is equal to the $\vec{0}$, i.e.,

$$b_1 s_1 + \cdots + b_{i-1} s_{i-1} + b_i v + b_{i+1} s_{i+1} + \cdots + b_n s_n = \vec{0}$$

Substituting the representation for v in terms of the vectors in S, we get

$$b_1 s_1 + \cdots + b_{i-1} s_{i-1} + b_i (a_1 s_1 + a_2 s_2 + \cdots + a_n s_n) + b_{i+1} s_{i+1} + \cdots + b_n s_n = \vec{0}$$

$$(b_1 + b_i a_1) s_1 + \cdots + (b_i a_i) s_i + \cdots + (b_n + b_i a_n) s_n = \vec{0}$$

Since the vectors in S are linearly independent, all the coefficients in the above equation are equal to 0. Thus, $b_i a_i = 0$ which implies $b_i = 0$ since we assumed $a_i \neq 0$. The above equation then becomes

$$b_1 s_1 + \cdots + b_{i-1} s_{i-1} + b_{i+1} s_{i+1} + \cdots + b_n s_n = \vec{0}$$

Since S is linearly independent, all the coefficients must equation 0. Thus, $b_j = 0, j = 1,2, \dots, n$ and S' is linearly independent.

Next, we need to show that S' spans the same space as S. Since we know S is a basis, it follows that $\text{span}(S') \subseteq \text{span}(S)$. If we can show $\text{span}(S) \subseteq \text{span}(S')$, we are done. Take any $u \in \text{span}(S)$, i.e,

$$u = c_1 s_1 + \cdots + c_i s_i + \cdots + c_n s_n$$

Given that $v = a_1 s_1 + a_2 s_2 + \cdots + a_n s_n$ and $a_i \neq 0$, we can write

$$s_i = -\frac{a_1}{a_i} s_1 - \cdots + \frac{1}{a_i} v - \cdots - \frac{a_n}{a_i} s_n$$

Substituting the expression for s_i into the expression for u, we have

$$u = c_1 s_1 + \cdots + c_i \left(-\frac{a_1}{a_i} s_1 - \cdots + \frac{1}{a_i} v - \cdots - \frac{a_n}{a_i} s_n \right) + \cdots + c_n s_n$$

$$= \left(c_1 - \frac{c_i a_1}{a_i} \right) s_1 + \cdots + \left(c_{i-1} - \frac{c_i a_{i-1}}{a_i} \right) s_{i-1} + \frac{c_i}{a_i} v + \left(c_{i+1} - \frac{c_i a_{i+1}}{a_i} \right) s_{i+1} + \cdots + \left(c_n - \frac{c_i a_n}{a_i} \right) s_n$$

Thus, any $u \in \text{span}\ (S)$ can be written as a linear combination of vectors in S' which implies $\text{span}(S) \subseteq \text{span}(S')$. ■

We are now in a position to prove our main theorem about the dimension of a vector space. This allows us to define the dimension of a vector space to be the number of elements in any of its bases.

Theorem 14. In a finite-dimensional vector space V, every basis has the same number of elements.

Proof: Let $S = \{s_1, s_2, \dots, s_n\}$ be a basis of V with a minimum number of elements. Let $T = \{t_1, t_2, \dots\}$ be any other basis of V. Since we assumed S has the minimum number of elements among the bases of V, we know that T has n or more elements. We need to prove that T necessarily has exactly n elements.

Since S is a basis, we can write t_1 as a linear combination of the vectors in S. Since $t_1 \neq \vec{0}$, at least one of the coefficients in the linear combination (say i_1) is not equal to zero. By Theorem 13, we can exchange t_1 with s_{i_1} and form the basis S_1 (consisting of t_1 and the other elements in S except for s_{i_1}).

Since S_1 is a basis, we can write t_2 as a linear combination of the vectors in S_1. Since $t_2 \neq \vec{0}$, at least one of the coefficients in the linear combination is not equal to zero. Further, it cannot be that only the coefficient of t_1 is non-zero, since that would imply that t_1 is dependent on t_2 which contradicts the assumption that T is linearly independent. So, the coefficient of at least one of the vectors from the original basis S must be non-zero (say i_2). By Theorem 13, we can exchange t_2 with s_{i_2} and form the basis S_2 (consisting of t_1, t_2 and the other elements in S except for s_{i_1} and s_{i_2}).

We continue the process until we arrive at the basis $S_n = \{t_1, t_2, \dots, t_n\}$. At this point, we must conclude that T cannot have more than these n vectors because any t_{n+1} that remains would be in the span of S_n (since it is a basis) and hence, t_{n+1} would be a linear combination of the other t_i vectors, contradicting that T is linearly independent. ■

The following theorems all rely, either directly or indirectly, on Theorem 14. Although not explicitly stated, the following theorems apply to finite dimensional vector spaces.

Theorem 15. A linearly independent set T of vectors in a vector space V cannot have more elements than the dimension of V.

Proof: This follows from the proof of Theorem 14 which only used the property that the vectors of T are linearly independent and not that T spans V. ■

Theorem 16. Any linearly independent S in a vector space V set can be expanded to form a basis for V.

Proof: If S also spans V, then we are done since S is a basis for V in that case.

Assume S does not span V. By Theorem 9, adding a vector to S that is not in the span of S will preserve linear independence. So, we just keep adding vectors until the resulting set spans V. By Theorem 15, the adding process necessarily stops after a finite number of steps. ■

Going in the other direction, a spanning set can be reduced to a basis.

Theorem 17. Any spanning set S of a vector space V can be reduced to a basis for V.

Proof:

Case 1: $S = \phi$ (the empty set). The empty set spans the zero vector space $\{\vec{0}\}$, because the span of the empty set is defined as $span(\phi) = \{\vec{0}\}$. The empty set is also linearly independent (vacuously, since there are no vectors to form a nontrivial linear combination). Thus, ϕ is already a basis for the zero vector space.

Case 2: $S = \{\vec{0}\}$. The set $\{\vec{0}\}$ spans $\{\vec{0}\}$, since $span(\{\vec{0}\}) = \{\vec{0}\}$. However, $\{\vec{0}\}$ is linearly dependent because $1\vec{0} = \vec{0}$ is a nontrivial linear combination yielding the zero vector. To reduce S to a basis, we remove $\vec{0}$, leaving the empty set ϕ, which is a basis for $\{\vec{0}\}$.

Case 3: S contains a non-zero vector, call it s_1. If $span(\{s_1\}) = V$, we are done. If not, then there $s_2 \in S$ such that $s_2 \notin span(\{s_1\})$. By Theorem 9, $span(\{s_1, s_2\})$ is linearly independent. If $span(\{s_1, s_2\}) = V$, we are done; otherwise, we continue adding vectors until we span V and thus form a basis of V using vectors from S. The process must stop since we assumed V is a finite dimensional vector space. ■

Theorem 18. In an n-dimensional space V, a set S (of n vectors) is linearly independent if and only if S spans V.

Proof: If S is a set of n linearly independent vectors, then (by Theorem 16), it can be expanded to be a basis for V. This implies that V has a basis with more than n vectors, which contradicts Theorem 14. Thus, no expansion of S is needed for S to span V.

If S spans V and is linearly dependent, then it can be reduced (by Theorem 17) to a basis for V. This implies that there is a basis for V with less than n vectors, which contradicts Theorem 14. Thus, if S spans V, it must be linearly independent. ■

2.6 Linear Systems

This section views linear systems from the perspective of vector spaces. The reader is assumed to be familiar with matrices and linear systems (including their solution using the Gaussian elimination [1]). For those needing to review these preliminaries, see Section 15 of Mathematical Thinking [2].

The **row space** of a matrix is the span of its rows. The **row rank** is the dimension of this space, i.e., the number of linearly independent rows.

For example, the row space of the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is all of $\mathbb{R}^2$.

The row space of $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$ is a 2-dimensional plane in $\mathbb{R}^2$ spanned by the vectors $(1,0,0)$ and $(0,1,0)$. The vector $(1,1,0)$ is redundant since it can be expressed as the sum of the other two row vectors.

Elementary row operations include (1) the exchange of two rows, (2) multiplication of a row by a scalar, and (3) adding the scalar multiple of one row to another.

Theorem 19. If matrix B is obtained from matrix A by a sequence of elementary row operations, then their row spaces are equal.

Proof: Suppose the rows of the matrix are $v_1, v_2, \dots, v_n$. If we switch the position of two rows to form matrix B, the span of the row vectors does not change. The same is true if we multiple a row of A by a scalar.

Concerning the third elementary row operation, suppose we change v_2 to the vector $v_2 + \alpha v_1$ where $\alpha \in \mathbb{R}$. The span of *B* (i.e., the matrix with row vectors $v_1, v_2 + \alpha v_1, v_3, \dots, v_n$) will include the vector $(v_2 + \alpha v_1) - \alpha v_1 = v_2$ which implies that $span(A) \subseteq span(B)$. Similarly, $v_2 + \alpha v_1 \in span(A)$ which implies $span(B) \subseteq span(A)$. Thus, $span(B) = span(A)$. ■

Theorem 20. The nonzero rows of a matrix in row echelon form [3] make up a linearly independent set.

Proof: Consider a matrix in row echelon form with the all-zero rows removed such as the example below.

$$\begin{pmatrix} 1 & m_{12} & m_{13} & m_{14} & \dots & \boldsymbol{m_{1j}} & \dots & m_{1n} \\ 0 & 0 & 1 & m_{24} & \dots & \boldsymbol{m_{2j}} & \dots & m_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \dots & m_{kn} \end{pmatrix}$$

Let v_i represent the i^{th} row in the above matrix. Suppose that a linear combination of the v_i vectors is equal to the zero vector, i.e.,

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = \vec{0}$$

We want to show that $a_i = 0, i = 1,2, \dots, k$ which will prove the row vectors are linearly independent.

The above equation can be expanded as

$$(a_1, a_1 m_{12}, a_1 m_{13} + a_2, a_1 m_{14} + a_2 m_{24} + \cdots + a_h, \dots,$$

$$\boldsymbol{a_1 m_{1j} + a_{2j} m_{2j} + \cdots + a_{k-1} m_{k-1,j} + a_k}, \dots)$$

Working from left to right, we first note that $a_1 = 0$. Plugging this into the second term, we have $0m_{12} = 0$ which does not yield any additional information. However, plugging $a_1 = 0$ into the third term, we see that $a_2 = 0$. Proceeding in this fashion, we have that $a_i = 0, i = 1,2, \dots, k$. For each column that has a 1 (that is also a leading 1 for a row), we get an equation that isolates one of the a_i and this is what allows us to progressively conclude each of the a_i coefficients is equal to 0.

While the position of the leading zero in each row will vary depending on the given matrix, the above procedure of deducing that all the coefficients are equal to zero is still valid. ■

In the language of vector spaces, the Gaussian elimination works by removing linearly dependent rows, while leaving the span of the row space unchanged. The elimination stops when no nontrivial linear relationships remain among the nonzero rows. In summary, the Gaussian elimination results in a basis for the row space of a matrix.

We can also define the **column space** of a matrix. It is the span of the set of its columns comprising a matrix. The **column rank** is the dimension of the column space, i.e., the number of linearly independent columns.

The transpose of a matrix A (written as A^T) is the result of interchanging its rows and columns, so that column i of matrix A is row i of A^T and vice versa. Further, it can be shown that $(AB)^T = B^T A^T$ for square matrices A and B.

Theorem 21. If matrix A is invertible, then A^T is also invertible and $\left(A^{-1}\right)^T = \left(A^T\right)^{-1}$.

Proof: Since A is invertible, $AA^{-1} = I \Rightarrow (AA^{-1})^T = (A^{-1})^T A^T = I$ and so, $(A^{-1})^T$ is a left inverse of A^T. Similarly, we can show that $(A^{-1})^T$ is a right inverse of A^T. Thus, $(A^{-1})^T$ is the inverse of A^T, i.e., $(A^{-1})^T = (A^T)^{-1}$. ■

Analogous to rows, elementary column operations (i.e., (1) exchanging columns, (2) multiplying a column by a scalar, or (3) multiplying a column by a scalar and adding to another column) do not change the column space.

Theorem 22. Row operations do not change the column rank.

Proof: This proof uses elementary matrices [4]. We also make use of the fact that a matrix times a column vector can be represented as a linear combination of the column vectors, e.g.,

$$\begin{pmatrix} 1 & -2 & 1 \\ 2 & 1 & 3 \\ 3 & 5 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 13 \\ 19 \end{pmatrix} \Rightarrow 1\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 2\begin{pmatrix} -2 \\ 1 \\ 5 \end{pmatrix} + 3\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 13 \\ 19 \end{pmatrix}$$

Further, we note that a linear dependence relation between some columns of a matrix A is equivalent to the existence of nonzero column vector v such that $Av = 0$. For example, columns 1 and 2 are multiples of each other in the 3×3 below.

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 6 & 7 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - 1\begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} + 0\begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Now to the proof …

First, note that the column rank of a matrix A is the size of the largest set of unrelated (i.e., independent) columns in A. If E is an elementary matrix and $B = EA$ (effectively a row operation on A) and $Av = 0$ for some nonzero column vector v then $Bv = EAv = 0$; conversely, if $Bv = 0$ then $Av = E^{-1}Bv = 0$. Thus, if a set of columns of A is linearly dependent then the **corresponding** set of columns of B is also linearly dependent, and vice versa. This means *A* and *B* have the same column rank.

The above proof is based on an answer to a question on Math Stack Exchange [5]. ■

Shown below are the steps in putting a matrix into reduced row echelon form. The final result gives us a basis for the row space, i.e., the rows with a 1 as the leading entry. Further, since the final result is in reduced echelon form, a basis for the column space is formed by the columns containing the leading entries. Thus, for a matrix in reduced echelon form, we can find bases for the row and column spaces by taking the rows or columns containing the leading entries, respectively.

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 2 & 2 & 4 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & -2 & 0 & -1 \\ 0 & 0 & 2 & 1 \\ 0 & -2 & 2 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 2 & 1 \\ 0 & -2 & 2 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{pmatrix} \rightarrow$$

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 0 & 1/2 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & -1/2 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We formalize the above concept in the following theorem.

Theorem 23. The row rank and column rank of a matrix are always equal.

Proof: If we put the matrix into reduced echelon form, then the row rank equals the number of rows with 1 as the leading entry since that equals the number of nonzero rows. Further, the number of leading entries equals the column rank because the set of columns containing leading entries consists of some of the vectors from the standard basis, and that set is linearly independent and spans the set of columns. So, in the reduced echelon form of a matrix, the row rank equals the column rank, because each equals the number of leading entries.

Theorem 19 and Theorem 22 prove that the row rank and column rank are not changed by using row operations in getting into reduced echelon form. So, the row rank and the column rank of the original matrix are also equal. ■

Given the above theorem, we can refer to the **rank of a matrix** without saying row or column rank since the two are equal.

Theorem 24. Given an $m \times n$ matrix A, the following statements are equivalent

- ***the rank of A is r***
- ***the solution space of $Ax = \vec{0}$ (known as the kernel or null space of A) has dimension $n - r$***

Proof: The following statements are equivalent:

- the rank of A is $r \Leftrightarrow$
- the Gaussian elimination applied to A results in r nonzero rows $\Leftrightarrow$
- the matrix in reduced echelon form that is row equivalent to A has r leading variables $\Leftrightarrow$
- when put in reduced echelon form the augmented matrix $(A|\vec{0})$ has $n - r$ free variables $\Leftrightarrow$
- the dimension of the null space of A is $n - r$. ■

As an example of the above theorem, consider a 4×7 matrix A which has the following reduced echelon form

$$A = \begin{bmatrix} 1 & a & b & 0 & c & 0 & e \\ 0 & 0 & 0 & 1 & d & 0 & f \\ 0 & 0 & 0 & 0 & 0 & 1 & g \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We can see that the rank of A is 3 and we are given that A has 7 columns. So, relative to the above theorem, we have that $r = 3$ and $n = 7$.

The theorem tells us that the kernel of A has rank $n - r = 4$. We can verify this for our example by solving the system $Ax = \vec{0}$ where $x^T = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)$.

The free variables are x_2, x_3, x_5, x_7.

From the first row, we can write x_1 in terms of the free variables, i.e.,

$$x_1 = -ax_2 - bx_3 - cx_5 - ex_7$$

From the second row, we can write x_4 in terms of the free variables, i.e.,

$$x_4 = -dx_5 - fx_7$$

From the third row, we can write x_6 in terms of the free variables, i.e.,

$$x_6 = -gx_7$$

So, the general solution to $Ax = \vec{0}$ is of the form

$$(-ax_2 - bx_3 - cx_5 - ex_7, x_2, x_3, -dx_5 - fx_7, x_5, -gx_7, x_7)$$

$$= (-ax_2, x_2, 0,0,0,0,0) + (-bx_3, 0, x_3, 0,0,0,0)$$

$$+(-cx_5, 0,0, -dx_5, x_5, 0,0) + (-\mathrm{e}x_7, 0,0, -fx_7, 0, -gx_7, x_7)$$

$$= x_2(-a, 1,0,0,0,0,0) + x_3(-b, 0,1,0,0,0,0) + x_5(-c, 0,0, -d, 1,0,0) + x_7(-e, 0,0, -f, 0, -g, 1)$$

So, the kernel of A, i.e., the solution space of $Ax = \vec{0}$, is of dimension 4 as it is spanned by the following set of linearly independent vectors.

$$\{(-a, 1,0,0,0,0,0), (-b, 0,1,0,0,0,0), (-c, 0,0, -d, 1,0,0), (-e, 0,0, -f, 0, -g, 1)\}$$

The following theorem summarizes the many of the preceding results in this section.

Theorem 25. Given an $n \times n$ matrix A, the following statements are equivalent:

- ***(1) the rank of A is n***
- ***(2) A is nonsingular, i.e., A^{-1} exists***
- ***(3) the rows of A form a linearly independent set***
- ***(4) the columns of A form a linearly independent set***
- ***(5) any linear system $Ax = b$ has one and only one solution, i.e., $x = A^{-1}b$.***

The above theorem and several variants of it are known as the Invertible Matrix Theorem. [6][7][8]

2.7 Isomorphisms

We have seen some examples that appear to be essentially the same in terms of the vector space properties, but have differences with respect to other (non-vector space aspects). For example, consider $\mathbb{R}^4$ and the set of all 2×2 real matrices $M_{2\times 2}$. Each is of dimension 4, addition is on a component basis, and scalar multiplication applies across each component. The relationship between these two vector spaces is known as an isomorphism.

Two vector spaces V and W are said to be isomorphic (written as $V \cong W$) if there exists a mapping $f: V \to W$ such that the following hold true:

- f is one-to-one (**injective**) and onto (**surjective**), i.e., f is **bijective**.
- $f(v_1 + v_2) = f(v_1) + f(v_2)$ for any $v_1, v_2 \in V$
- $f(av) = af(v)$ for $a \in \mathbb{R}, v \in V$

The latter two properties can be combined into one, i.e.,

$$f(av_1 + bv_2) = af(v_1) + bf(v_2), \qquad a, b \in \mathbb{R},\ v_1, v_2 \in V$$

Example 1. Going back to our example, we can define a mapping between $\mathbb{R}^4$ and $M_{2\times 2}$ as follows:

For $v = (v_1, v_2, v_3, v_4) \in \mathbb{R}^4$, define $f: \mathbb{R}^4 \to M_{2\times 2}$ by $f(v) = \begin{pmatrix} v_1 & v_2 \\ v_3 & v_4 \end{pmatrix} \in M_{2\times 2}$.

To show that f is injective, we need to show the $v \neq u$ implies that $f(v) \neq f(u)$, or conversely, $f(v) = f(u)$ implies that $u = v$. We will use the latter approach here. So, assume $u = (u_1, u_2, u_3, u_4)$ and $v = (v_1, v_2, v_3, v_4)$, and that $f(v) = f(u)$ which implies

$$\begin{pmatrix} v_1 & v_2 \\ v_3 & v_4 \end{pmatrix} = \begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix}$$

Using the definition of matrix equality, we have that $v_i = u_i, i = 1,2,3,4$ which implies $v = u$.

Clearly, f is surjective since for every element in $M_{2\times 2}$ there is an element in $\mathbb{R}^4$ that is mapped to it.

With u and v as defined above, we have

$$f(au + bv) = \begin{pmatrix} au_1 + bv_1 & au_2 + bv_2 \\ au_3 + bv_3 & au_4 + bv_4 \end{pmatrix} = a\begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix} + b\begin{pmatrix} v_1 & v_2 \\ v_3 & v_4 \end{pmatrix} = af(u) + bf(v)$$

Exercise 1. Show that the set of real polynomials of degree 2 or less, i.e., $\{a_0 + a_1 x + a_2 x^2 : a_0, a_1, a_2 \in \mathbb{R}\}$ is isomorphic to $\mathbb{R}^3$.

An isomorphism between a vector space and itself is known as an **automorphism**.

Example 2. The vector space $\mathbb{R}$ is automorphic to $\mathbb{R}$ under the mapping $f(x) = cx$ where $c \in \mathbb{R}$ is a constant.

Theorem 26. An isomorphism maps the zero vector of one vector space to the zero vector of the other vector space.

Proof: Let f be an isomorphism between vector spaces V and W. For any $v \in V$, we have $f(\vec{0}_V) = f(0v) = 0f(v) = \vec{0}_W \in W$. ■

Theorem 27. The inverse of an isomorphism is an isomorphism.

Proof: The inverse of a bijective function is (by definition) also a bijective function.

Let f be an isomorphism from vector space V to W. Take any $w_1, w_2 \in W$ where $f(v_1) = w_1$ and $f(v_2) = w_2$. Concerning the linearity property, we have that

$$f^{-1}(aw_1 + bw_2) = f^{-1}\big(af(v_1) + bf(v_2)\big) = f^{-1}\big(f(av_1 + bv_2)\big)$$
$$= av_1 + bv_2 = af^{-1}(w_1) + bf^{-1}(w_2)$$

Thus, f^{-1} is an isomorphism. ■

. . .

An **equivalence relation** R is a binary relation on the members of a set S that is reflexive ($aRa\ \forall a \in S$), symmetric (if aRb then bRa) and transitive ($aRb, bRc \Rightarrow aRc$). A common example is the integers modulo some integer n. An equivalence relation partitions the associated set into equivalence classes.

For example, the integers modulo 7 partition the integers into seven equivalence classes (one class for each remainder when an integer is divided by 7).

For example, $[6] = \{\dots, -15, -8, -1, 6, 13, 20, \dots\}$ is the equivalence class of all integers whose remainder is 6 when divided by 7, e.g., $20 = 2 \cdot 7 + 6,\ -15 = -3 \cdot 7 + 6$.

Theorem 28. Isomorphism is an equivalence relation between vector spaces.

Proof: In terms of the reflexive property, every vector space V is isomorphic to itself. To see this, consider the identity map $id: V \to V$ where $id(v) = v\ \forall v \in V$. Clearly, id is bijective. Further, the linearity property holds for the identity mapping, i.e.,

$$id(av + bu) = av + bu = a\ id(v) + b\ id(u)$$

In terms of the symmetric property, assume there exists an isomorphism f from vector space V to W. By Theorem 27, f^{-1} is an isomorphism from W to V.

In terms of transitivity, assume f is an isomorphism from vector space V to W, and g is an isomorphism from W to X. The composition of g with f (i.e., $g \circ f: V \to X$) is a bijection from V to X since the composition of bijections is a bijection. [9]

Concerning the linearity property, we have that

$$g \circ f(av_1 + bv_2) = g\big(f(av_1 + bv_2)\big) = g\big(af(v_1) + bf(v_2)\big)$$
$$= ag\big(f(v_1)\big) + bg\big(f(v_2)\big) = a(g \circ f)(v_1) + b(g \circ f)(v_2)$$

Thus, we have established the transitivity property. ■

Example 1 and Exercise 1 suggest that there may be one equivalence class per dimension. As we shall see, that is the case.

Theorem 29. Isomorphic finite dimensional vector spaces have the same dimension.

Proof: The approach here is to demonstrate a mapping between the bases of two isomorphic vector spaces V and W.

Let f be an isomorphism from V to W, and let $S = \{s_1, s_2, \dots, s_n\}$ be a basis for V (implying V is of dimension n). The image of S in W is $T = \{f(s_1), f(s_2), \dots, f(s_n)\}$. We want to show that T is a basis for W, and thus, conclude that W is also of dimension n.

Take any $w \in W$. Since f is an isomorphism, it is an onto mapping and so, there exists $v \in V$ such that $f(v) = w$. Writing v in terms of the basis S, we have

$$w = f(v) = f(a_1 s_1 + a_2 s_2 + \cdots + a_n s_n) = a_1 f(s_1) + a_2 f(s_2) + \cdots + a_n f(s_n)$$

Thus, T spans W.

We also need to establish the linear independence of the vectors in set T. Consider a linear combination of the vectors in T that equal the zero vector, i.e.,

$$\vec{0}_w = b_1 f(s_1) + b_2 f(s_2) + \cdots + b_n f(s_n) = f(b_1 s_1 + b_2 s_2 + \cdots + b_n s_n)$$

Since f is a one-to-one mapping, we know that it maps $\vec{0}_v$ to one and only one vector, i.e., $\vec{0}_w$. So, $b_1 s_1 + b_2 s_2 + \cdots + b_n s_n = \vec{0}_v$ and thus, $b_1 = b_2 = \cdots = b_n = 0$. ■

The converse of Theorem 29 is also true.

Theorem 30. Finite dimensional vector spaces of the same dimension are isomorphic.

Proof: Our approach will be to show any vector space V of dimension n is isomorphic to $\mathbb{R}^n$. It then follows by transitivity (from Theorem 28) that all such spaces are isomorphic to each other.

Let $S = \{s_1, s_2, \dots, s_n\}$ be a basis for an n-dimensional vector space V. Define a mapping $f: V \to \mathbb{R}^n$ using the representation function, i.e.,

$$f(v) = Rep_S(v) = (a_1, a_2, \dots, a_n)_S \in \mathbb{R}^n, v \in V$$

Since v has a unique representation in terms of the vectors of S, $f(v)$ is mapped to a unique element of $\mathbb{R}^n$. Thus, f is a well-defined function.

To show that f is injective (one-to-one), take $v = a_1 s_1 + \cdots + a_n s_n$ and $u = b_1 s_1 + \cdots + b_n s_n$. If $f(v) = f(u)$, then $Rep_S(v) = (a_1, \dots, a_n)_S = (b_1, \dots, b_n)_S = Rep_S(u)$ which implies that $a_i = b_i, i = 1,2, \dots, n$. Thus, $v = a_1 s_1 + \cdots + a_n s_n = b_1 s_1 + \cdots + b_n s_n = u$.

For every $w = (w_1, \dots, w_n) \in \mathbb{R}^n$ there is a $v = w_1 s_1 + \cdots + w_n s_n$ that is mapped to w by f. So, f is onto (surjective).

Lastly, we need to demonstrate linearity. Take $v, u \in V$ as defined above.

$$f(xv + yu)$$

$$= Rep_S(xv + yu) = Rep_S\big((xa_1 + yb_1)s_1 + \cdots + (xa_n + yb_n)s_n\big)$$

$$= (\, xa_1 + yb_1, \dots, xa_n + yb_n) = x(a_1, \dots, a_n) + y(b_1, \dots, b_n) = xRep_S(v) + yRep_S(u)$$

$$= xf(v) + yf(u)$$

Thus, $f = Rep_S$ is an isomorphism, and every n-dimensional space is isomorphic to $\mathbb{R}^n$. ■

The above two theorems imply the following theorem.

Theorem 31. Each finite-dimensional vector space is isomorphic to one and only one of the R^n.

2.8 Homomorphisms

When two spaces are of different dimensions, it is not possible to define a bijection between them, e.g., $\mathbb{R}^3$ and $\mathbb{R}^2$. In such cases, it is still possible to define what is called a **homomorphism**. A homomorphism requires only the linearity property, i.e., a function f that maps from vector space V to vector space W is a homomorphism if

$$f(av + bu) = af(v) + bf(u), \qquad a, b \in \mathbb{R}, \qquad u, v \in V$$

The above condition is equivalent to the following:

$$f(v + bu) = f(v) + bf(u), \qquad b \in \mathbb{R}, \qquad u, v \in V$$

A mapping between vector spaces that has the above linearity property is also known as a **linear transformation**.

In the context of vector spaces, the terms homomorphism, linear transformation, linear map and linear mapping all refer to the same type of thing.

Example 1. Define $f: \mathbb{R}^3 \to \mathbb{R}^2$ as $f(x, y, z) = (x + y, y + z)$. To show linearity and thus, prove f is a homomorphism, take $(x, y, z), (x', y', z') \in \mathbb{R}^3$. We have that

$$f\big(a(x, y, z) + b(x', y', z')\big) = f\big((ax + bx'), (ay + by'), (az + bz')\big)$$

$$= \big((ax + bx') + (ay + by'), (ay + by') + (az + bz')\big)$$

$$= (ax + ay + bx' + by', ay + az + by' + bz')$$

$$= (ax + ay, ay + az) + (\, bx' + by', by' + bz')$$

$$= a(x + y, y + z) + b(x' + y', y' + z') = af(x, y, z) + bf(x', y', z')$$

In the above, we have made use of the definition of vector addition and scalar multiplication in the vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$.

The function f maps $\mathbb{R}^3$ to all of $\mathbb{R}^2$, i.e., f is surjective (onto). To see this, take any $(z_1, z_2) \in \mathbb{R}^2$ and note that $f(z_1, 0, z_2) = (z_1, z_2)$.

We have that $f(1,2,1) = (3,3)$ and $f(2,1,2) = (3,3)$. So, the f^{-1} is not a function and thus, f is not an isomorphism.

Example 2. We can also go from smaller to larger, e.g., $f: \mathbb{R} \to \mathbb{R}^2$ as defined by $f(x) = (x, 2x)$. In terms of linearity, we have

$$f(ax + by) = (ax + by, 2ax + 2by) = (ax, 2ax) + (by, 2by)$$

$$= a(x, 2x) + b(y, 2y) = af(x) + bf(y)$$

Example 3: Given vectors space V and W, one can always define the zero homomorphism, i.e., $f: V \to W$ such that $f(v) = \vec{0}_W, \forall v \in V$.

A homomorphism maps the zero vector to the zero vector. The same proof as Theorem 26 applies here. This fact can be helpful in identifying when a function is not a homomorphism (or not an isomorphism). For example, consider the mapping $f: \mathbb{R}^3 \to \mathbb{R}^2$ defined by $f(x, y, z) = (x + 2, y + z)$. We can quickly see this is not a homomorphism since $f(0,0,0) = (2,0)$.

A homomorphism $f: V \to W$ is determined by the mapping of the basis for V.

Let V and W be vector spaces and $S = \{s_1, \ldots, s_n\}$ be a basis for V. A function $f: S \to W$ is said to be **extended linearly** to a function $f': V \to W$, if $\forall v \in V, f'(v) = a_1 f(s_1) + \cdots + a_n f(s_n)$ where $v = a_1 s_1 + \cdots + a_n s_n$.

Theorem 32. If $S = \{s_1, \ldots, s_n\}$ is a basis for vector space V, and $T = \{t_1, \ldots, t_n\}$ is a multiset of elements from vector space W, then there exists a unique homomorphism from V to W that maps s_i to t_i for $i = 1, 2, \ldots, n$.

(A **multiset** is a collection of elements with the possibility of repeated elements as opposed to sets which do not allow repeated elements.)

Proof: Define a function $g: S \to W$ such that $g(s_i) = t_i, i = 1,2, \ldots, n$. Linearly extend g to a function $f: V \to W$ such that for every $v = a_1 s_1 + \cdots + a_n s_n \in V, f(v) = a_1 t_1 + \cdots + a_n t_n$. This function is well-defined since each $v \in V$ is uniquely represented in terms of the vectors in S.

(A **well-defined function** is one in which each input is assigned to exactly one output.)

To show linearity, take $v = a_1 s_1 + \cdots + a_n s_n$ and $u = b_1 s_1 + \cdots + b_n s_n$, and consider

$$f(av + bu) = f\big(a(a_1 s_1 + \cdots + a_n s_n) + b(b_1 s_1 + \cdots + b_n s_n)\big)$$

$$= f\big((aa_1 + bb_1)s_1 + \cdots + (aa_n + bb_n)s_n\big)$$

$$= (aa_1 + bb_1)t_1 + \cdots + (aa_n + bb_n)t_n$$

$$= a(a_1 t_1 + \cdots + a_n t_n) + b(b_1 t_1 + \cdots + b_n t_n)$$

$$= af(v) + bf(u)$$

In terms of uniqueness, assume there is another homomorphism h such that $h(s_i) = t_i, i = 1,2, \ldots, n$. For any $v = a_1 s_1 + \cdots + a_n s_n \in V$, we have

$$h(v) = a_1 h(s_1) + \cdots + a_n h(s_n) = a_1 t_1 + \cdots + a_n t_n$$

$$= a_1 f(s_1) + \cdots + a_n f(s_n) = f(v)$$

Thus, h and f are identical since the have the same value for every $v \in V$. ■

Why do we use a multiset in the previous theorem?

The following examples will provide some insight.

Let $V = \mathbb{R}^2$ with standard basis $S = \{e_1, e_2\}$, and let $W = \mathbb{R}$. Define T to be the multiset $\{1,1\}$. The theorem guarantees a unique linear map $f: \mathbb{R}^2 \to \mathbb{R}$ such that $f(e_1) = 1$ and $f(e_2) = 1$. This map is $f(x, y) = x + y$.

If the theorem was restricted to "regular" sets (i.e., not multisets), we would exclude the valid solution in the above example.

As a second example, consider $f: \mathbb{R}^5 \to \mathbb{R}^3$ where we take S to be the standard basis for $\mathbb{R}^5$ and $T = \{e_1, e_2, e_3, \vec{0}, \vec{0}\}$. In this case, we map the first three elements of the standard basis for $\mathbb{R}^5$ to the standard basis for $\mathbb{R}^3$, and we map e_4 and e_5 to the zero vector in $\mathbb{R}^3$. The mapping that fulfills requirements is $f(x_1, x_2, x_3, x_4, x_5) = x_1e_1 + x_2e_2 + x_3e_3$.

Exercise: Show that the set of all function $\boldsymbol{F(V, W)}$ from vector space $\boldsymbol{V}$ to vector space $\boldsymbol{W}$ is a vector space where addition and scalar multiplication are defined as follows:

- $(\boldsymbol{f_1} + \boldsymbol{f_2})(\boldsymbol{v}) = \boldsymbol{f_1}(\boldsymbol{v}) + \boldsymbol{f_2}(\boldsymbol{v}), \quad \boldsymbol{f_1}, \boldsymbol{f_2} \in \boldsymbol{F(V, W)}$
- $(\boldsymbol{\alpha f})(\boldsymbol{v}) = \boldsymbol{\alpha f}(\boldsymbol{v}), \quad \boldsymbol{f} \in \boldsymbol{\mathcal{L}(V, W)}, \quad \boldsymbol{\alpha} \in \mathbb{R}$

The set of all homomorphisms from vector space V to vector space W is denoted by $\mathcal{L}(V, W)$

Theorem 33. The set $\mathcal{L}(\boldsymbol{V}, \boldsymbol{W})$ is a vector space where addition and scalar multiplication are defined the same as for $\boldsymbol{F(V, W)}$.

Proof: $\mathcal{L}(V, W)$ is non-empty since it contains the zero homomorphism, i.e., $f(v) = \vec{0}_W, \forall v \in V$. Since $\mathcal{L}(\mathrm{V}, \mathrm{W})$ is a non-empty subset of the vector space of all functions from V to W, we only to show that $\mathcal{L}(\mathrm{V}, \mathrm{W})$ is closed under addition and scalar multiplication (this follows from Theorem 2). We leave this as an exercise for the reader. ■

For a function $f: X \to Y$, X is known as the domain, Y is known as the codomain, and $f(X) = \{f(v): v \in V\} \subseteq Y$ is the image (sometime "range") of X under f. The dimension of the image space is known as the **function's rank**.

Omitting the onto (surjection) requirement does not make an essential difference between isomorphisms and homomorphisms, since each homomorphism is onto some space, i.e., its image. However, omitting the one-to-one (injection) condition does make a difference, since a homomorphism may map several elements of the domain to a single element in the codomain. We saw that in Example 1 where, for example, $f(1,2,1) = f(2,1,2) = (3,3)$.

Theorem 34. Under a homomorphism, the image of any non-empty subspace of the domain is a subspace of the codomain.

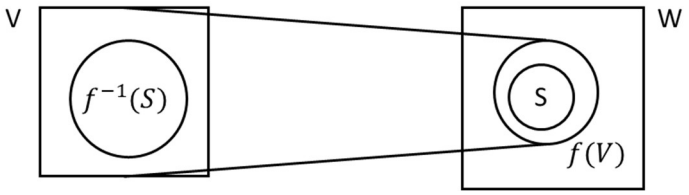
Proof: Let $f: V \to W$ be homomorphism and let S be a subspace of V. The image of S, i.e., $f(S)$, is a subset of the codomain W, which is not empty because S is not empty. So, we only need to show that $f(S)$ is closed under the linear combinations of two vectors (by Theorem 2).

Take $f(u), f(v) \in f(S)$ and $a, b \in \mathbb{R}$, we have that $af(u) + bf(v) = f(au) + f(bv) = f(au + bv) \in f(S)$ since $au + bv \in S$. ■

The following theorem is the reverse of the previous theorem.

Theorem 35. For a homomorphism $\boldsymbol{f}: \boldsymbol{V} \to \boldsymbol{W}$, the inverse image of a subspace $\boldsymbol{S}$ in the range $\boldsymbol{f(V)}$ is a subspace of the domain $\boldsymbol{V}$.

Proof: S is a subspace, and so, it contains the zero vector $\vec{0}_W$. Since f is a homomorphism, we have that $f(\vec{0}_V) = \vec{0}_W$. Thus, $f^{-1}(S)$ is not empty, and by Theorem 2, we only need to show that $f^{-1}(S)$ is closed under the linear combinations of two vectors in $f^{-1}(S)$.



Take $u, v \in f^{-1}(S), a, b \in \mathbb{R}$ which implies $f(u), f(v) \in S$. Further, $f(au + bv) = af(u) + bf(v) \in S$ since S is a subspace of $f(V)$. This, in turn, implies that, $au + bv \in f^{-1}(S)$. ■

The **null space or kernel** of a homomorphism $f: V \to W$ is the inverse mapping of $\vec{0}_W$. The dimension of the null space is known as the function's nullity. In terms of notation, we write

$$\mathcal{N}(f) = f^{-1}(\vec{0}_W) = \{v \in V: f(v) = \vec{0}_W\}$$

Example 4. Consider the function $f: \mathbb{R}^3 \to \mathbb{R}^2$ as $f(x, y, z) = (x + y, y + z)$ which we previously proved was a homomorphism. The elements of $\mathbb{R}^3$ that are mapped to $(0,0)$ are all $(x, y, z) \in \mathbb{R}^3$ such that $x = -y$ and $z = -y$. So, the null space of f consists of all vectors of the form

$$(-y, y, -y) = y(-1,1,-1)$$

So, the nullity of f is 1.

All the vectors in the image space of f can be written in the following form

$$(x + y, y + z) = (x, 0) + (y, y) + (0, z) = x(1,0) + y(1,1) + z(0,1)$$

So, we can map to any $(s, t) \in \mathbb{R}^2$ by setting $x = s, y = 0$ and $z = t$. Thus, the rank of the image space of f is 2. Notice that the rank of the image space plus the nullity equals the dimension of the domain $\mathbb{R}^3$. As we shall prove in the next theorem, this is always true for a homomorphism.

Theorem 36. The sum of a homomorphism's rank and its nullity equals the dimension of its domain.

Proof: Let $f: V \to W$ be a homomorphism and let $S_{\mathcal{N}} = \{s_1, s_2, \dots, s_k\}$ be a basis for the null space of f, i.e., the nullity is k. Assume the dimension of V is $n \geq k$. By Theorem 16, we can expand $S_{\mathcal{N}}$ to a basis for all of V, i.e., $S_V = \{s_1, \dots, s_k, s_{k+1}, \dots, s_n\}$. If we can show that $T = \{f(s_{k+1}), \dots, f(s_n)\}$ is a basis for the image of f, then we are done.

We first show that T is linearly independent. Take $a_{k+1}f(s_{k+1}) + \cdots + a_n f(s_n) = \vec{0}_W$. This implies that $f(a_{k+1}s_{k+1} + \cdots + a_n s_n) = \vec{0}_W$ and so, $a_{k+1}s_{k+1} + \cdots + a_n s_n$ is in the null space of f. Since $S_{\mathcal{N}}$ is a basis for the null space, there exist scalars $a_1, \dots, a_k$ such that

$$a_1 s_1 + \cdots + a_k s_k = a_{k+1}s_{k+1} + \cdots + a_n s_n$$

$$\Rightarrow a_1 s_1 + \cdots + a_k s_k - a_{k+1}s_{k+1} - \cdots - a_n s_n = \vec{0}_V$$

However, the vectors in S_V are linearly independent since S_V is a basis for V. Thus, $a_i = 0, i = 1,2, \dots, n$ which implies that T is linearly independent.

To show that T spans the image of f, take any member of the image, call it $f(v)$, and represent v as a linear combination of the vectors in S_V, i.e.,

$$v = a_1 s_1 + \cdots + a_n s_n$$

We then have the following

$$f(v) = f(a_1 s_1 + \cdots + a_n s_n) = a_1 f(s_1) + \cdots + a_k f(s_k) + a_{k+1} f(s_{k+1}) + \cdots + a_n f(s_n)$$

Since $s_i, i = 1,2, \ldots, k$ are in the null space, $f(s_i) = 0, i = 1,2, \ldots, k$. So,

$$f(v) = a_{k+1} f(s_{k+1}) + \cdots + a_n f(s_n) \in \text{span}(T)$$

Thus, T spans the entire image of f.

So, if the nullity of f is k, then the dimension of the image of f (i.e., the rank of f) is $n - k$, and their sum is the dimension of V. ■

Exercise 1. Prove that for a homomorphism, the image of a linearly dependent set is linearly dependent.

Theorem 37. For an n-dimensional vector space V and homomorphism $f: V \to W$, the following statements are equivalent

(1) f is one-to-one (injective).

(2) f has an inverse from its range $f(V)$ back to V and the inverse is a homomorphism.

(3) $\mathcal{N}(f) = \{\vec{0}_V\}$, i.e., the nullity of f is 0.

(4) The rank of f is n.

(5) If $\{s_1, s_2, \ldots, s_n\}$ is a basis for V then $\{f(s_1), f(s_2), \ldots, f(s_n)\}$ is a basis $f(V)$.

Proof: The plan is to show $(1) \Leftrightarrow (2)$, and then show that $(1) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5) \Rightarrow (2)$.

Assume (1) is true, i.e., f is one-to-one. This implies that the existence of $f^{-1}: f(V) \to V$. Take a linear combination of elements from $f(V)$, i.e., $af(u) + bf(v)$. Next, apply f^{-1} to the linear combination:

$$f^{-1}\big(af(u) + bf(v)\big) = f^{-1}\big(f(au + bv)\big) = au + bv = af^{-1}\big(f(u)\big) + bf^{-1}\big(f(v)\big)$$

So, if a homomorphism (f in this case) has an inverse then the inverse must also be a homomorphism.

Assume (2) is true. In general, if a function f has an inverse f^{-1}, then f^{-1} must be one-to-one. So, (1) holds true in this case.

$(1) \Rightarrow (3)$ holds true since all homomorphisms map the zero vector to the zero vector, and a one-to-one map sends each member of the domain to one member of the codomain.

$(3) \Rightarrow (4)$ holds true since (by definition) the rank plus the nullity equals the dimension of the domain.

To establish $(4) \Rightarrow (5)$, we only need to show that $\{f(s_1), f(s_2), \dots, f(s_n)\}$ spans $f(V)$ since by (4), the dimension of $f(V)$ is n (this follows from Theorem 18). So, take any $f(v) \in f(V)$ and write v in terms of $\{s_1, s_2, \dots, s_n\}$, i.e., $v = a_1 s_1 + \cdots + a_n s_n$. We have that

$$f(v) = f(a_1 s_1 + \cdots + a_n s_n) = a_1 f(s_1) + \cdots + a_n f(s_n)$$

Finally, we need to prove that $(5) \Rightarrow (2)$. Assume that (5) is true. Take $u, v \in V$ such that $u \neq v$. The two vectors have unique representations in terms of the basis for V, i.e.,

$$u = a_1 s_1 + \cdots + a_n s_n, \qquad v = b_1 s_1 + \cdots + b_n s_n$$

At least one pair of the coefficients are not equal, say $a_j \neq b_j$. The assumption that $\{f(s_1), f(s_2), \dots, f(s_n)\}$ is a basis for $f(V)$ and $a_j \neq b_j$ implies

$$f(u) = a_1 f(s_1) + \cdots + a_n f(s_n) \neq b_1 f(s_1) + \cdots + b_n f(s_n) = f(v)$$

Thus, f is one-to-one. ■

2.9 Representing a Homomorphism with a Matrix

A homomorphism (linear transformation) can be represented by a matrix. Will first look at a simple example and then extend to the general case.

Notation: In what follows, we use column vectors to represent a vector in terms of a basis. Further, when not using the standard basis, we will mark the representation vector with a subscript that indicates the name of the basis; otherwise (when using the standard basis), no subscript will be used (thereby indicating use of the standard basis).

Example 1. Let $L: \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x + y \\ 3x - y \end{bmatrix}$$

Assume the standard basis for $\mathbb{R}^2$ is used for both the domain and codomain. We leave it as an exercise for the reader to verify that L is a linear transformation.

This transformation can be represented using matrix multiplication, i.e.,

$$L\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = A\begin{bmatrix} x \\ y \end{bmatrix}, A = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}$$

For example, let's evaluate $\begin{bmatrix} -2 \\ 2 \end{bmatrix}$ in two different ways; first using the initial definition of the linear transformation in this example and then using the matrix approach.

$$L\left(\begin{bmatrix} -2 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 2(-2) + 2 \\ 3(-2) - 2 \end{bmatrix} = \begin{bmatrix} -2 \\ -8 \end{bmatrix}$$

$$L\left(\begin{bmatrix} -2 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 + 2 \\ -6 - 2 \end{bmatrix} = \begin{bmatrix} -2 \\ -8 \end{bmatrix}$$

Example 2. This is the same as the previous example, except that we use non-standard bases for the domain and codomain. For the domain, we will use the basis

$$S = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$$

For the codomain, we will use the basis

$$T = \left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\}$$

First, we compute the mapping of each basis element under the transformation L, i.e.,

$$L\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = \begin{bmatrix}3\\2\end{bmatrix}, \qquad L\left(\begin{bmatrix}1\\-1\end{bmatrix}\right) = \begin{bmatrix}1\\4\end{bmatrix}$$

Next, we express the above results in terms of the basis for the codomain, i.e.,

$$Rep_T\left(\begin{bmatrix}3\\2\end{bmatrix}\right) = \begin{bmatrix}1\\2\end{bmatrix}_T, \qquad Rep_T\left(\begin{bmatrix}1\\4\end{bmatrix}\right) = \begin{bmatrix}-3\\4\end{bmatrix}_T$$

Finally, we use the above vectors to form the transformation matrix, i.e.,

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

As a check, let's evaluate $\begin{bmatrix}-2\\2\end{bmatrix}$ in two different ways. We first need to covert $\begin{bmatrix}-2\\2\end{bmatrix}$ to its representation under basis S, i.e., $\begin{bmatrix}0\\-2\end{bmatrix}_S$.

$$L\left(\begin{bmatrix}-2\\2\end{bmatrix}\right) = \begin{bmatrix}-2\\-8\end{bmatrix} = \begin{bmatrix}6\\-8\end{bmatrix}_T$$

$$L\left(\begin{bmatrix}-2\\2\end{bmatrix}\right) = \begin{bmatrix}1 & -3\\2 & 4\end{bmatrix}\begin{bmatrix}0\\-2\end{bmatrix}_S = \begin{bmatrix}6\\-8\end{bmatrix}_T$$

Let's formalize the notation for the concepts discussed in the previous examples.

Let V be vector space of dimension n with basis S, and W a vector space of dimension m with basis T. Let f be a homomorphism (linear transformation) from V to W. Next, we represent the mapping (by f) of each basis element of S in terms of T as follows:

$$Rep_T(f(s_1)) = \begin{bmatrix}a_{11}\\a_{21}\\\dots\\a_{m,1}\end{bmatrix}_T, \dots, Rep_T(f(s_n)) = \begin{bmatrix}a_{1,\mathrm{n}}\\a_{21}\\\dots\\a_{m,1}\end{bmatrix}_T$$

If we put the above column vectors together, we form what is known as the matrix representation of f with respect to the bases S and T, as shown below.

$$Rep_{S,T}(f) = \begin{bmatrix}a_{11} & a_{12} & \dots & a_{1,n}\\a_{21} & a_{22} & \dots & a_{2,n}\\\dots & \dots & \dots & \dots\\a_{m,1} & a_{m,2} & \dots & a_{m,n}\end{bmatrix}_{S,T}$$

Theorem 38. If $A = Rep_{S,T}(f)$ is the matrix representation of a homomorphism (linear transformation) from vector space V (of dimension n with basis S) to vector space W (of dimension m with basis T), then the mapping of a vector $v \in V$ by f is given by the product $Av_S = w_T$ where v is represented as a column vector in terms of the basis S and the output is a vector in W expressed as column vector in terms of the basis T.

Proof: Take any $v \in V$ and express it in terms of the basis $S = \{s_1, s_2, \dots, s_n\}$, i.e.,

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix}_S$$

Next, we represent the mapping of each vector in S is terms of $T = \{t_1, t_2, \dots, t_m\}$, i.e.,

$$f(s_1) = a_{11}t_1 + a_{21}t_2 + \cdots + a_{m,1}t_m$$

$$f(s_2) = a_{12}t_1 + a_{22}t_2 + \cdots + a_{m,2}t_m$$

…

$$f(s_n) = a_{1,n}t_1 + a_{2,n}t_2 + \cdots + a_{m,n}t_m$$

Expanding $f(v)$ in terms of T, we have

$$w_1t_1 + w_2t_2 + \cdots + w_mt_m = f(v) = f(v_1s_1 + v_2s_2 + \cdots + v_ns_n)$$

$$= v_1f(s_1) + v_2f(s_2) + \cdots + v_nf(s_n)$$

$$= v_1\left(a_{11}t_1 + a_{21}t_2 + \cdots + a_{m,1}t_m\right) + v_2(a_{12}t_1 + a_{22}t_2 + \cdots + a_{m,2}t_m) + \cdots$$

$$+v_n(a_{1,n}t_1 + a_{2,n}t_2 + \cdots + a_{m,n}t_m)$$

Equating like components from both sides of the above, we have

$$w_1 = a_{11}v_1 + a_{12}v_2 + \cdots + a_{1,n}v_n$$

$$w_2 = a_{21}v_1 + a_{22}v_2 + \cdots + a_{2,n}v_n$$

…

$$w_m = a_{m,1}v_1 + a_{m,2}v_2 + \cdots + a_{m,n}v_n$$

In term of matrices, the above can be written as

$$\begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_m \end{bmatrix}_{\mathrm{T}} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1,n} \\ a_{21} & a_{22} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix}_S$$

The above is of the form $w_T = Av_S$ as required by the statement of the theorem. ■

In the above proof, notice that column i of matrix A is $f(s_i)$ represented in basis T. This gives a methodology for constructing the matrix that represents a particular linear transformation.

We can also go in the opposite direction, i.e., for a given matrix, and associated homomorphism between appropriate vector spaces can be specified.

Theorem 39. A homomorphism between vector spaces of appropriate dimensions, with respect to any pair of bases, can be defined for any given matrix A.

Proof: Assume that A is of dimension $m \times n$. In this case, the dimension of the domain (call it V) needs to be n and the dimension of the codomain (call it W) needs to be m. Let S be a basis for the domain, and T be a basis for the codomain. Define the mapping $f: V \to W$ as

$$f(v) = Rep_T(w) = A \cdot Rep_S(v)$$

Take $u, v \in V, a, b \in \mathbb{R}$ and consider the following:

$$f(au + bv) = A\big(aRep_S(u) + bRep_S(v)\big) = aA \cdot Rep_S(u) + bA \cdot Rep_S(v) = af(u) + bf(v)$$

This show that f is a linear transformation. ■

The above theorem still holds even if there is a linear dependency among the rows (or columns) of the matrix A. The key idea is that the matrix represents a linear transformation between vector spaces of appropriate dimensions, regardless of whether the matrix is of full rank or not.

Clarifying the Role of Rank and Linear Dependence:

- Linear Dependence in Rows or Columns:
    - If the rows of A are linearly dependent, it means the matrix does not have full row rank (i.e., $rank(A) < m$, where m is the number of rows).
    - Similarly, if the columns are linearly dependent, the matrix does not have full column rank (i.e., $rank(A) < n$, where n is the number of columns).
- Effect on the Linear Transformation: The rank of A determines the dimension of the image (range) of the linear transformation f represented by A.
    - If the rows are dependent, the image of f is a subspace of W with dimension $rank(A)$, which is less than m.
    - If the columns are dependent, the null space of f is nontrivial, i.e.,
    $$\mathcal{N}_f = n - rank(A) > 0$$
- However, f is still a well-defined linear transformation from V to W, even if it is not surjective (due to row dependencies) or injective (due to column dependencies).

Several of the statements in the above require proof which we will address in the theorems that follow.

Example 3. Given the matrix $A = \begin{bmatrix} 4 & 1 \\ -2 & 3 \end{bmatrix}$, define an associated homomorphism f.

We first note that the two vector spaces need to be of dimension 2 given that A is a 2×2 matrix. Secondly, we need to select a basis for the domain V and a basis for the codomain W.

Let $S = \left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\1\end{bmatrix}\right\}$ be the basis for V, and $T = \left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\2\end{bmatrix}\right\}$ be the basis for W.

Define our homomorphism as follows:

$$f(v) = Rep_T(w) = \begin{bmatrix} 4 & 1 \\ -2 & 3 \end{bmatrix} \cdot Rep_S(v)$$

For example, take the vector $v = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \in V$. To apply our homomorphism to v, we first need to determine $Res_S(v)$.

$$v = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -\frac{4}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \frac{3}{5}\begin{bmatrix} 3 \\ 1 \end{bmatrix} \Rightarrow Res_S(v) = \frac{1}{5}\begin{bmatrix} -4 \\ 3 \end{bmatrix}_S$$

We then have

$$\begin{bmatrix} 4 & 1 \\ -2 & 3 \end{bmatrix} \cdot Rep_S(v) = \begin{bmatrix} 4 & 1 \\ -2 & 3 \end{bmatrix} \cdot \frac{1}{5}\begin{bmatrix} -4 \\ 3 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -13 \\ 17 \end{bmatrix}_T = Rep_T(w)$$

In terms of the standard basis for $\mathbb{R}^2$, we have that

$$w = \frac{1}{5}\left(-13\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 17\begin{bmatrix} 0 \\ 2 \end{bmatrix}\right) = \frac{1}{5}\begin{bmatrix} -4 \\ 21 \end{bmatrix}$$

If we used the standard bases for V and W, then conversions would not be needed for the input or the output.

Example 4. Given the matrix $A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$, determine the associated homomorphism f using the standard basis for both the domain V and the codomain W.

The desired homomorphism is defined as follows. The notation v^T is the transpose of vector v. The transpose of a vector is a column vector. This is needed for the multiplication in the definition below.

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = f(v) = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} v^T, v = (v_1, v_2) \in V$$

Since $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} v^T = \begin{bmatrix} v_1 + v_2 \\ 0 \end{bmatrix}$, we have that $w_1 = v_1 + v_2, w_2 = 0$ which we can write as

$$f\left(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right) = \begin{bmatrix} v_1 + v_2 \\ 0 \end{bmatrix}$$

Thus, the dimension of the image of f (i.e., its rank) is 1.

Exercise 1. Find the null space of f in the above example.

In general, we have the following theorem concerning the rank of a homomorphism and any matrix that represents it.

Theorem 40. The rank of a matrix equals the rank of any linear transformation that it represents.

Proof: See Theorem 2.3 in "Linear Algebra/Any Matrix Represents a Linear Map" [10]. ■

Theorem 41. Let $f: V \to W$ be a linear transformation represented by a matrix A. f is onto (surjective) if and only if the rank of A equals the number of its rows. f is one-to-one (injective) if and only if the rank of A equals the number of its columns.

In the second part of the above theorem, f could be one-to-one with a subspace of the codomain.

Proof: Concerning the surjective part of the theorem, the dimension of $f(V)$ (i.e., the rank of f) equals the rank of A by Theorem 40. The dimension of the codomain of f (which is the vector space W) equals the number of rows in A. So, if the rank of A equals the number of its rows, then the dimension of the $f(V)$ equals the dimension of the codomain. Thus, $f(V) = W$ and f is onto.

Concerning the injective part of the theorem, a linear transformation is one-to-one if and only if it is an isomorphism between its domain and its image $f(V)$ if and only if $\dim(V) = \dim(f(V))$. However, the number of columns in A is the dimension of V (i.e., the domain of f) and by Theorem 40, the rank of A equals the dimension of $f(V)$. ■

Example 5: Consider the following matrix as a representation of a linear transformation $f: \mathbb{R}^3 \to \mathbb{R}^4$ using the standard bases for both the domain and codomain.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

For $v = (x, y, z) \in \mathbb{R}^3$, we have the following mapping by f into $\mathbb{R}^4$.

$$f(v) = Av^T = \begin{bmatrix} x \\ 0 \\ z \\ y \end{bmatrix}$$

The rank of A is 3 but the number of rows in A is 4. So, by the above theorem, the mapping is not onto. For example, there is no pre-image for $(0,1,0,0)$.

However, the rank of A does equal the number of columns in A, and in fact, the mapping is one-to-one with a subspace of $\mathbb{R}^4$.

If instead of A (in the above example) we used $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, then the mapping would be

$$f(v) = Bv^T = \begin{bmatrix} x \\ 0 \\ 0 \\ y \end{bmatrix}$$

As before, the mapping is not onto.

The mapping is also not one-to-one since the rank of B is 2 but B has 3 columns. For example, $(1,2,3)$ and $(1,2,5)$ both get mapped to $(1,0,0,2)$. In fact, $(1,0,0,2)$ has an infinite number of pre-images of the form $(1,2,z)$.

A one-to-one and onto (bijective) linear transformation is said to be **nonsingular**; otherwise, it is said to be singular, i.e., a linear transformation is nonsingular if and only if it is an isomorphism.

A square matrix A is **nonsingular** if it is the matrix of coefficients of a homogeneous system (i.e., $A\vec{x} = 0$) with a unique solution, i.e., A^{-1} exists; otherwise, a matrix is said to be singular.

Theorem 42. (1) A nonsingular linear transformation $f: V \to W$ is represented by a square matrix. (2) A square matrix represents a nonsingular linear transformation if and only if it is a nonsingular matrix (i.e., has an inverse). Thus, a matrix represents an isomorphism if and only if it is square and nonsingular.

Proof: (1) Assume that f is nonsingular and of rank n. By Theorem 40, the rank of any matrix A representing f equals the rank of f. By Theorem 41 and the fact that f is bijective, we have that the number of rows of A and the number of columns of A each equal the rank of f. So, A is an $n \times n$ matrix.

(2) A square $n \times n$ matrix A is nonsingular if and only if its row rank is n, which is true if and only if A is of rank is n (by Theorem 23 and the definition of rank for a matrix), which is true if and only if the rank of f is n (by Theorem 40), i.e., $f(V)$ is of dimension n, which is true if and only if f is an isomorphism (by Theorem 29). The last statement holds true because the domain of f (i.e., V) is n-dimensional since it equals the number of columns in A. ■

For example, any linear transformation represented by $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 4 \end{bmatrix}$ is nonsingular since A is nonsingular.

Any linear transformation represented by $B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 0 \\ 4 & 2 & 0 \end{bmatrix}$ is singular since B is singular.

2.10 Change of Basis

Consider an isomorphism $i: V \to V$ defined by $i(v) = v, \forall v \in V$ (the identity mapping). Let S be the basis for the domain and T be the basis for the codomain. By Theorem 38, the matrix $A = Rep_{S,T}(i)$ effectively performs a change of basis for each vector in V, i.e., $v_T = Av_S$ (putting the vectors in column form).

Theorem 43. To convert from the representation of a vector v in basis S to basis T multiple v_S on the left by the matrix $A = Rep_{S,T}(i)$ where i is the identity mapping. In addition, if $Mv_S = v_T$ then M is the change of basis matrix.

Proof: We proved the first part of the theorem in the paragraph preceding this theorem.

For the second part of the theorem, we note that with respect to bases S and T, the matrix M represents a linear transformation whose action is to map each vector to itself (but as expressed in a different basis), and therefore, $M = Rep_{S,T}(i) = A$. ■

Example 1. We are given two bases for $\mathbb{R}^3$, i.e.,

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \qquad T = \left\{ \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \right\}$$

Our task is to find a matrix A that allows us to convert a vector written in terms of S to the same vector written in terms of T.

Let $f: \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $i(v) = v$. Let S be the basis for the domain and T be the basis for the codomain. Next, we compute $i(s_j) = s_j$ for each $s_j \in S$ in terms of the basis T.

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}_S = \begin{bmatrix} .5 \\ 0 \\ 0 \end{bmatrix}_T, \qquad \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}_S = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}_T, \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_S = \begin{bmatrix} 0 \\ 0 \\ .5 \end{bmatrix}_T$$

The above results are the columns for our change of matrix $A = Res_{S,T}(f)$, i.e.,

$$A = \begin{bmatrix} .5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & .5 \end{bmatrix}$$

Let's test this with the vector $v_S = \begin{bmatrix} 1 \\ .5 \\ 3 \end{bmatrix}_S = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + .5 \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$

$$Av_S = \begin{bmatrix} .5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & .5 \end{bmatrix} \begin{bmatrix} 1 \\ .5 \\ 3 \end{bmatrix}_S = \begin{bmatrix} .5 \\ 1 \\ 1.5 \end{bmatrix}_T = v_T$$

Expressing v_T in terms of the standard basis, we have

$$\begin{bmatrix} .5 \\ 1 \\ 1.5 \end{bmatrix}_T = .5 \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + (1.5) \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$$

So, when expressed in terms of the standard basis, v_S and v_T yield the same vector, as expected.

Theorem 44. A matrix effects a change of basis if and only if it is nonsingular.

Proof: See Lemma 1.4 in "Linear Algebra/Changing Representations of Vectors" [11]. ■

Theorem 45. A matrix is nonsingular if and only if it represents the identity map with respect to some pair of bases.

Proof: This follows from Theorem 43 and Theorem 44. ■

. . .

It is possible to convert from a representation of a homomorphism $f: V_{S_1} \to W_{T_1}$ to $f: V_{S_2} \to W_{T_2}$ where the subscript on each vector space indicates its basis. As shown in Figure 1, there are two ways to determine a matrix representing function f that maps a vector v (represented in basis S_2 for vector space V) to a vector w (represented in basis T_2 for vector space W):

- We can directly compute $Res_{S_2,T_2}(f)$.
- We can compute and then apply $Rep_{S_2,S_1}(i), Rep_{S_1,T_1}(f), Rep_{T_1,T_1}(i)$ where i is the identity function.

Thus, we have the equation

$$Res_{S_2,T_2}(f) = Rep_{T_1,T_2}(i) \cdot Rep_{S_1,T_1}(f) \cdot Rep_{S_2,S_1}(i)$$

Application of the formula is via multiplication on the left, i.e.,

$$Res_{S_2,T_2}(f)\boldsymbol{v_{S_2}} = Rep_{T_1,T_2}(i) \cdot Rep_{S_1,T_1}(f) \cdot Rep_{S_2,S_1}(i)\boldsymbol{v_{S_2}} = \boldsymbol{w_{T_2}}$$

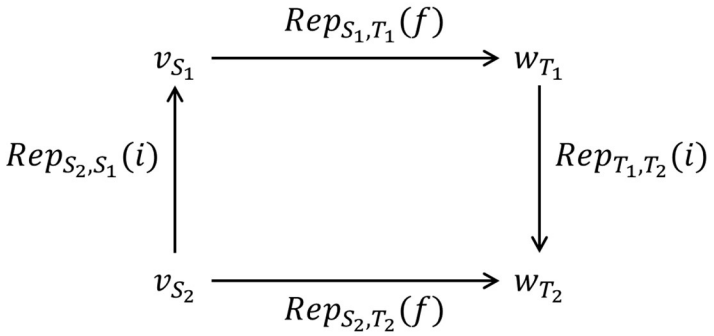


Figure 1. Basis change for a homomorphism

We have proven the following theorem.

Theorem 46. To convert from the matrix $\boldsymbol{Rep_{S_1,T_1}(f)}$ representing a homomorphism $\boldsymbol{f}$ with respect to bases $\boldsymbol{S_1}$ and $\boldsymbol{T_1}$ to the matrix $\boldsymbol{Res_{S_2,T_2}(f)}$ representing $\boldsymbol{f}$ with respect to bases $\boldsymbol{S_2}$ and $\boldsymbol{T_2}$, one can use the following formula:

$$\boldsymbol{Res_{S_2,T_2}(f) = Rep_{T_1,T_2}(i) \cdot Rep_{S_1,T_1}(f) \cdot Rep_{S_2,S_1}(i)}$$

Example 2. Before we start our example, we need a little background in complex analysis and the rotation of vectors.

[Background] In the complex plane, numbers (effectively 2-dimensional vectors) can be represented as polar coordinates using something known as Euler's formula [12], i.e., $re^{i\theta} = r(\cos\theta + i\sin\theta)$ where r is the length of the vector and θ is the angle the vector makes with the real axis (assuming the vector's tail is at the origin). For the discussion here, assume $r = 1$. To rotate a vector by an angle α, we simply multiply by $e^{i\alpha}$. So, given vector $v = e^{i\theta} = \cos\theta + i\sin\theta$, the vector

$$e^{i\alpha}e^{i\theta} = e^{i(\alpha+\theta)} = \cos(\theta+\alpha) + i\sin(\theta+\alpha)$$

$$= (\cos\theta\cos\alpha - \sin\theta\sin\alpha) + i(\sin\alpha\cos\theta + \sin\theta\cos\alpha)$$

The second line in the above comes from the sum of angles formulas for the sine and cosine. The above formula can be represented in matrix notation, i.e.,

$$Av = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}\begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix} = \begin{bmatrix}\cos\theta\cos\alpha - \sin\theta\sin\alpha \\ \sin\alpha\cos\theta + \sin\theta\cos\alpha\end{bmatrix}$$

The matrix A is a θ radian counterclockwise rotation of a vector in $\mathbb{R}^2$ (which is isomorphic to the complex plane) when using the standard basis for both the domain and codomain**. [End of Background]**

Now to our example, let $f: \mathbb{R}^2 \to \mathbb{R}^2$ be the homomorphism defined by a $\frac{\pi}{2}$ counterclockwise rotation. From the background discussion, we know this rotation can be accomplished via multiplication on the left by the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Note that A was defined in terms of the standard basis for the domain and the codomain, i.e., $\mathcal{E}_2$.

Our task is to find the matrix that represents f if we use $S = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ as the basis for the domain and $T = \left\{\begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$ as the basis for the codomain. We will find the solutions in two different ways, i.e., directly by computing $Rep_{S,T}(f)$ and via the longer route $Rep_{\mathcal{E}_2,T} \cdot A \cdot Rep_{S,\mathcal{E}_2}(f)$.

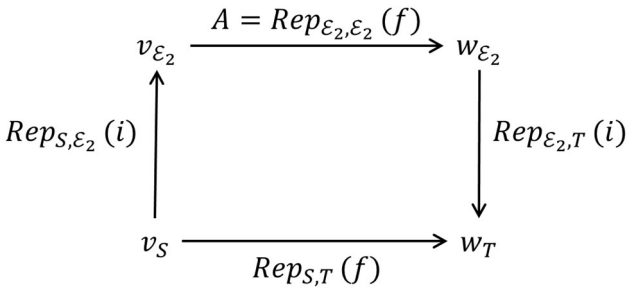


Figure 2. Change of basis example

$Rep_{S,\mathcal{E}_2}(i) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ since the ending basis is $\mathcal{E}_2$ and we can directly use the vectors of S as the columns. We already have A. For the last leg of the journey, we could write $\mathcal{E}_2$ in terms of the basis T, or we could determine $\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}^{-1}$. Taking the latter approach, we have $Rep_{\mathcal{E}_2,T} = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & .5 \\ 1 & 0 \end{bmatrix}$. So, using this approach gives us

$$Rep_{\mathcal{E}_2,T} \cdot A \cdot Rep_{S,\mathcal{E}_2}(f) = \begin{bmatrix} 0 & .5 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} .5 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} .5 & 0 \\ -1 & -1 \end{bmatrix}$$

Taking the other approach, we need to represent each vector in S in terms of T.

$$f\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = (.5)\begin{bmatrix} 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$f\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = A\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} = 0 \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Thus, $Rep_{S,T}(f) = \begin{bmatrix} .5 & 0 \\ -1 & -1 \end{bmatrix}$ which matches the result obtained using the other approach.

As an additional check, take the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{E}_2}$, write it in terms of the basis S, multiply it by $Rep_{S,T}(f)$ and then convert it back to $\mathcal{E}_2$, i.e.,

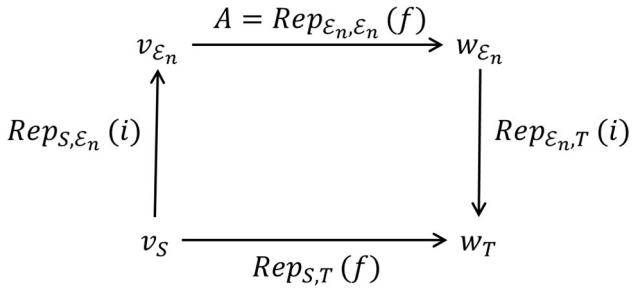
$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{E}_2} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_S, \qquad \begin{bmatrix} .5 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}_S = \begin{bmatrix} .5 \\ -1 \end{bmatrix}_T = \begin{bmatrix} -1 \\ 1 \end{bmatrix}_{\mathcal{E}_2}$$

This is the same result as applying A to $\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{E}_2}$, i.e.,

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{E}_2} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}_{\mathcal{E}_2}$$

. . .

Example 3. In the situation shown below, assume that matrix A is known, and we want to find the matrix representation for the homomorphism f with the indicated change of bases, i.e., basis S for the domain and basis T for the codomain.



The path $Rep_{S,\mathcal{E}_n}(i) \to Rep_{\mathcal{E}_n,\mathcal{E}_n}(f) \to Rep_{\mathcal{E}_n,T}(i)$ requires much less computation than computing $Rep_{S,T}(f)$. If $n = 10$, $Rep_{S,\mathcal{E}_n}(i)$ is formed using the vectors from S, we already known A and (as we saw in Example 2) and $Rep_{\mathcal{E}_n,T}(i)$ requires the inversion of the $n \times n$ matrix $Rep_{T,\mathcal{E}_n}(i)$. On the other hand, to compute $Rep_{S,T}(f)$, we need to solve 10 systems of equations (each with 10 unknowns).

. . .

Rectangular $m \times n$ matrices A and B are said to be equivalent if $B = PAQ$ for an $m \times m$ invertible matrix P and $n \times n$ invertible matrix Q. **Equivalent matrices** [13] represent the same linear transformation $f: V \to W$ under two different choices of basis pairs for vector spaces V and W, with P and Q being the change of basis matrices for W and V, respectively.

To prove the above statement, let

- $f: V \to W$ be a linear transformation from vector space V to vector space W
- $S = \{s_1, \dots, s_n\}$ and $S' = \{s'_1, \dots, s'_n\}$ be bases for V
- $T = \{t_1, \dots, t_n\}$ and $T' = \{t'_1, \dots, t'_n\}$ be bases for W
- $A = Res_{S,T}(f)$ and $B = Res_{S',T'}(f)$
- Q be the change of basis matrix from S' to S, i.e., $Q = Res_{S',S}(i)$. So, for any $v \in V$, we have that $v_S = Qv_{S'}$

- P be the change of basis matrix from T to T', i.e., $P = Res_{T,T'}(i)$. So, for any $w \in W$, we have that $w_{T'} = Pw_T$.

Putting the above together, we have (by Theorem 46)

$$B = Res_{S',T'}(f) = Res_{T,T'}(i)\, Res_{S,T}(f) Res_{S',S}(i) = PAQ$$

. . .

Matrix equivalence is an equivalence relation on the set of rectangular matrices.

The equivalence of two matrices of the same size can be characterized by the following conditions:

- The matrices can be transformed into one another by a combination of elementary row and column operations.
- Two matrices are equivalent if and only if they have the same rank.

Theorem 47. Any $m \times n$ matrix of rank k is matrix equivalent to the $m \times n$ matrix that is all zeroes except that the first k diagonal entries are ones.

Proof: The matrix described in the statement of the proof is shown in the figure below where I is the $k \times k$ identity matrix, and the Z entries are all zero matrices.

$$\begin{bmatrix} I & Z \\ Z & Z \end{bmatrix}$$

In terms of the proof, we first reduce the given matrix by a series of elementary row operations. Represent the row operations as multiplication by elementary matrices on the left and collect these elementary matrices to form P. Next, use the leading entries to do column reduction and finish by swapping the columns to put the leading ones along the diagonal. Represent the column operations as multiplication by elementary matrices on the right, and collect these elementary matrices to form Q. ■

Theorem 48. Matrix equivalent classes are characterized by their rank, i.e., two matrices of the same size are matrix equivalent if and only if they have the same rank.

Example 4. To illustrate Theorem 47, consider the matrix $A = \begin{bmatrix} 1 & 1 & 2 & 0 \\ 2 & 3 & 5 & 3 \\ 1 & 2 & 2 & 1 \\ 1 & 1 & 3 & 1 \end{bmatrix}$. We first row reduce matrix A to determine the matrix P that was mentioned in the proof. The right-hand side of the calculations keep track of the elementary row operations. Matrix P is shown in bold at the final step of the calculations.

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

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

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

We could have continued with the row reduction which would have resulted in the matrix $Q = I$. To make the result a bit more interesting, we finished the reduction with an elementary column operation. The matrix Q is shown in bold.

$$\left[\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & \mathbf{1} & \mathbf{-1} & \mathbf{0} & \mathbf{0} \\ 0 & 1 & 0 & 0 & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ 0 & 0 & 1 & 0 & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ 0 & 0 & 0 & 1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{array}\right]$$

Putting the above results together, we have the following

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

Since we started with an invertible square matrix, the final result was I rather than $\begin{bmatrix} I & Z \\ Z & Z \end{bmatrix}$. In other words, the rows and columns of zeros were not needed in this case.

2.11 Projections

Given $u = (x_1, x_2, \dots, x_n), v = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$, we define the **dot product** of u with v as

$$u \cdot v = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \in \mathbb{R}$$

For an n-dimensional vector space V other than $\mathbb{R}^n$, we can use the representation of the vectors in V to define the dot product. Two vectors are orthogonal if their dot product is zero.

Given an n-dimensional vector space V, the vectors in the set $M = \{v_1, v_2, \dots, v_k\} \subset V$ are said to be **mutually orthogonal** if the dot product of the representation of each pair of vectors from M is zero, i.e., $Rep_S v_i \cdot Rep_S v_j = 0$ for $1 \leq i, j \leq k$ where S is a basis for V.

The **orthogonal projection** of a vector v onto the line spanned by the vector u is defined to be the vector

$$proj_u(v) = \frac{u \cdot v}{u \cdot u}\, u$$

Note that the expression $\frac{v \cdot u}{u \cdot u}$ is a scalar.

To motivate the above definition of "vector projection", consider the 2-dimensional case depicted in Figure 3. The vector αu is the projection of v onto u. In order to find α, we solve the equation

$$u \cdot (v - \alpha u) = 0 \;\Rightarrow\; u \cdot v - \alpha\, u \cdot u = 0 \;\Rightarrow\; \alpha = \frac{u \cdot v}{u \cdot u} \;\Rightarrow\; proj_u(v) = \frac{u \cdot v}{u \cdot u}\, u$$

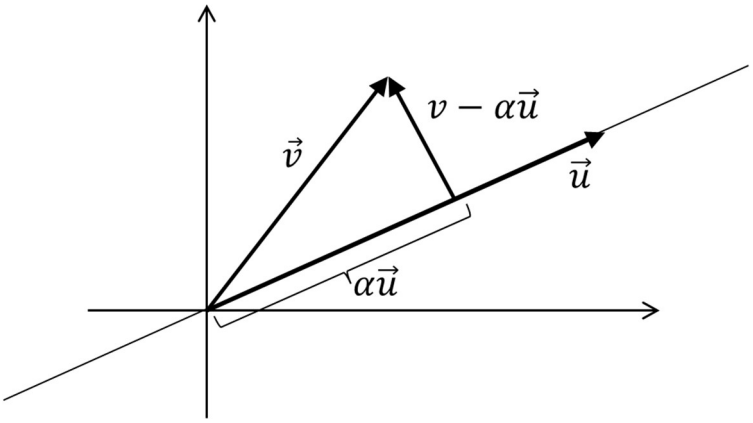


Figure 3. Projection in 2 dimensions

Theorem 49. Any set of non-zero and mutually orthogonal vectors $\{v_1, v_2, \dots, v_k\}$ in an n-dimensional vector space V are linearly independent.

Proof: Consider any linear combination of the vectors that equals zero, i.e., $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$. Taking the dot product of v_i with both sides of the equation, we get

$$a_1v_i \cdot v_1 + a_2v_i \cdot v_2 + \cdots + a_nv_i \cdot v_n = v_i \cdot 0 = 0$$

Since the vectors are mutually orthogonal, all the terms on the left are zero, except for $a_iv_i \cdot v_i$. This implies the $a_i = 0$. ■

Theorem 50. In an n-dimensional vector space V, if the vectors in a set S of size n are mutually orthogonal and nonzero then the set is a basis for the space.

Proof: As we saw from the previous theorem, the vectors in S are linearly independent, and by Theorem 18, S forms a basis for V. ■

A basis for a vector space consisting of mutually orthogonal vectors is referred to as an **orthogonal basis**.

A **normed vector space** is a vector space V over which a norm (length) function has been defined. In terms of notation, we write the norm of $v \in V$ as $\|v\|$. The norm maps each vector to a positive real number, and satisfies the following conditions:

1. Non-negativity: $\forall v \in V, \|v\| \geq 0$.
2. Positive definiteness: $\|v\| = 0$ if and only if $v = \vec{0}$.
3. Absolute homogeneity: $\forall \alpha \in \mathbb{R}, v \in V$ we have that $\|\alpha v\| = |\alpha| \, \|v\|$.
4. Triangle inequality: $\forall u, v \in V, \|u + v\| \leq \|u\| + \|v\|$.

For example, given $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$, we can define the norm $\|v\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$. The complex numbers in n dimensions (denoted $\mathbb{C}^n$) are a normed vector space. $\mathbb{R}^n$ and $\mathbb{C}^n$ are also normed vector spaces under the l^p norms for $1 \leq p \leq \infty$.

For a real number $p \geq 1$, the p-norm or l^p-norm of vector $v = (v_1, \dots, v_n)$ is defined by

$$\|v\|_p = (|v_1|^p + \cdots + |v_n|^p)^{\frac{1}{p}}$$

The l^∞-norm or maximum norm (or uniform norm) is the limit of the l^p norms as $p \to \infty$ and is given by

$$\|v\|_\infty = \max\{|v_1|, \ldots, |v_n|\}$$

In the above formulas, the expression $|v_i|$ is the absolute value of v_i in the case of $\mathbb{R}^n$. For $\mathbb{C}^n$, $|z_i|$ is the modulus of the complex number $z = x + iy$, i.e., $|z| = \sqrt{x^2 + y^2}$.

Given a mutually orthogonal set of vectors that form a basis for a vector space, we can normalize each vector by dividing it by its norm (length) to get what is known as an **orthonormal basis**. For example, the standard basis for $\mathbb{R}^n$ is an orthonormal basis.

The **Gram–Schmidt orthogonalization process** [14] is an algorithm for orthogonalizing a set of vectors. Given k linearly independent vectors $v_1, v_2, \ldots, v_k$, the Gram-Schmidt process generates a set of mutually orthogonal vectors (the vectors $u_1, u_2, \ldots, u_k$) which can be normalized (the vectors $w_1, w_2, \ldots, w_k$) as shown below.

$$u_1 = v_1, \qquad w_1 = \frac{u_1}{\|u_1\|}$$

$$u_2 = v_2 - proj_{u_1}(v_2), \qquad w_2 = u_2/\|u_2\|$$

$$u_3 = v_3 - proj_{u_1}(v_3) - proj_{u_2}(v_3), \qquad w_3 = u_3/\|u_3\|$$

$$u_4 = v_4 - proj_{u_1}(v_4) - proj_{u_2}(v_4) - proj_{u_3}(v_4), \qquad w_4 = u_4/\|u_4\|$$

$$\ldots$$

$$u_k = v_k - \sum_{i=1}^{k-1} proj_{u_i}(v_i)\,, \qquad w_k = \frac{w_k}{\|w_k\|}$$

2.12 Determinants

Recall from Theorem 25 that there exist multiple equivalent conditions for a square matrix to be invertible. In this section, we discuss yet another equivalent condition concerning matrix invertibility. The condition is related to something known as the determinant of a square matrix. We go directly to the definition of a determinant. If the reader is interested in the motivation for this definition, see "Linear Algebra/Exploration" [15].

A **determinant** is a real-valued function defined on square matrices that has the following properties:

(1) The determinant of the identity matrix is 1.

(2) The exchange of two rows multiplies the determinant by -1.

(3) Multiplying a row by a number multiplies the determinant by that number.

(4) Adding a multiple of one row to another row does not change the determinant.

The above statements hold true if we replace "row" by "column".

In terms of notation, we write the determinant of a square matrix A as $\det A$.

Theorem 51. The following properties hold true for determinants:

(1) A matrix with two identical rows has a determinant of zero.

(2) A matrix with a zero row has a determinant of zero.

(3) A matrix is nonsingular if and only if its determinant is nonzero.

(4) ***The determinant of a matrix in row echelon form is the product of the elements along the diagonal of the matrix.***

Proof: This theorem also holds true if we replace "row" with "column".

(1) When two rows are exchanged, the sign of the determinant changes (per the definition) but the matrix remains unchanged (given that the two rows are identical) and thus, the determinant before and after the switch must have the same value. This can only happen if the value of the determinant is zero.

(2) Multiply the zero row by any real number $a \neq 0$. By definition, this changes the determinant by a multiple of a, but it also leaves the row unchanged, and thus, the value of the determinant remains unchanged. This can only happen if the value of the determinant is zero.

Property (4) has two cases, i.e., the given matrix A is singular or nonsingular.

If the matrix A is singular, then it has a zero row (as we are given that it is in row echelon form). By property (2) of this theorem, the determinant of A is zero. So, in the case of a singular matrix in row echelon form, the matrix's determinant equals the product of its diagonal elements.

If the matrix A is nonsingular and in row echelon form, then none of its diagonal entries is zero. Let's say the diagonal entries are $a_{ii} \neq 0,\ i = 1,2, \dots, n$ (assuming A is $n \times n$). Using item (3) from the definition of a determinant, we can divide out the diagonal entries and get $a_{11}a_{22} \dots a_{nn}$ times the determinant of a matrix (call it A') that has the value 1 for each of its diagonal elements. Using item (4) from the definition of a determinant, A' can be reduced to the identity matrix which has determinant 1 (by definition). Thus, $\det A = a_{11}a_{22} \dots a_{nn}$, i.e., the $\det A$ equals the product of its diagonal elements.

(3) If A is nonsingular, then we can put A into row echelon form with non-zero elements along the main diagonal by using properties (3) and (4) from the definition of a determinant. The determinant of the resulting matrix A' is non-zero by part (4) of this theorem. Thus, $\det A = a \det A' \neq 0$ for some scalar a.

If A is singular, then it has a zero row when put into row echelon form A'. By property (2) of this theorem, $\det A' = 0$. By properties (2), (3) and (4) from the definition of a determinant, $\det A$ and differs from $\det A'$ by a scalar multiple, and thus, $\det A = 0$. ■

Using the definition of "determinant" and some properties from the previous theorem gives us a way to compute the determinant of a matrix.

Example 1. Given $A = \begin{bmatrix} 2 & 4 & 6 \\ 1 & 3 & 5 \\ 5 & 2 & 4 \end{bmatrix}$, find its determinant. In what follows, we use the notation $|A|$ to represent the determinant of a matrix. In the following computation, we first factor a 2 from the first row and then use row operations to get the matrix in row echelon form. The last step is to multiply the diagonal elements.

$$\det A = \begin{vmatrix} 2 & 4 & 6 \\ 1 & 3 & 5 \\ 5 & 2 & 4 \end{vmatrix} = 2\begin{vmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 5 & 2 & 4 \end{vmatrix} = 2\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -8 & -11 \end{vmatrix} = 2\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{vmatrix} = 2(5) = 10$$

In the definition of "determinant", we left open the possibility that multiple functions could fulfill the determinant requirements. The following theorem settles this issue.

Theorem 52. For each $n \in \mathbb{N}$, if there is an $n \times n$ determinant function then it is unique.

Proof: Assume there are two functions f and g that map from the set of $n \times n$ matrices to $\mathbb{R}$ and that satisfies the conditions of the definition of "determinant" and Theorem 51. Given an $n \times n$ matrix A, fix some way of performing row reductions to get the matrix into row echelon form, keeping track of row-scaling factors and how the sign alternates on row swaps. Lastly, multiply the diagonal elements in the row echelon form of the matrix times any scaling factors to calculate the determinant. This process only uses the definition of "determinant" and Theorem 51. Therefore, f and g must return the same value. Since f and g produce the same output on every input, they must be the same function. ■

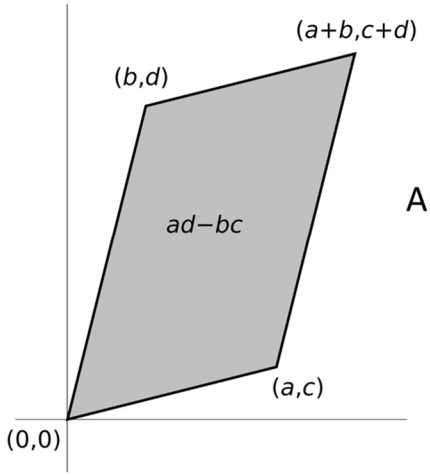
There are several ways to define the determinant of a square matrix. The **Leibnitz formula** [16] provides a definition of "determinant" using permutations. If A is an $n \times n$ matrix, where a_{ij} is the entry in row i and column j, then

$$\det A = \sum_{\tau \in S_n} sgn(\tau) \prod_{i=1}^{n} a_{i\tau(i)} = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$$

The function $sgn(\tau)$ returns value 1 if the permutation τ is even and the value -1 if the permutation is odd. S_n is the set of permutations on the integers $1,2, \ldots, n$. This formula defines the determinate as the sum of all products of matrix elements, where each product includes one element from each row and a different column. Using the Leibnitz formula, it is possible to prove that for each $n \in \mathbb{N}$ there exists an $n \times n$ determinant function. [17]

The determinant of the transpose of A (i.e., A^T) equals the determinant of A, i.e., $\det(A^T) = \det(A)$. This can be proven by inspecting the Leibniz formula.

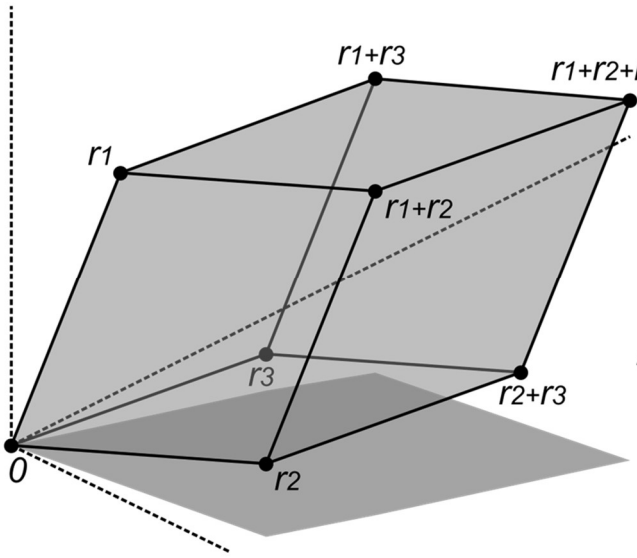
There is a geometric interpretation of determinants. For example, consider the vectors $(a, c), (b, d) \in \mathbb{R}^2$ and the parallelogram that they form (as shown in Figure 4). The area of the parallelogram is the determinant of the matrix formed by the two vectors (you need to take the absolute value if the determinant is negative).



$$\text{Area} = \det \begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc$$

Figure 4. Determinant as the area of a parallelogram

The same pattern holds in three dimensions. Figure 5 depicts the parallelepiped formed by the vectors $r_1, r_2, r_3 \in \mathbb{R}^3$. The volume of the parallelepiped is the determinant of the matrix formed by the three vectors (you need to take the absolute value if the determinant is negative).



$$r_1 = (x_1, y_1, z_1)$$
$$r_2 = (x_2, y_2, z_2)$$
$$r_3 = (x_3, y_3, z_3)$$

$$\text{Area} = \det \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

Figure 5. Determinant as the area of a parallelepiped

In $\mathbb{R}^n$ the box (or parallelepiped) formed by the linearly independent vectors $v_1, v_2, \ldots, v_n$ is defined to be the set

$$\{a_1 v_1 + a_2 v_2 + \cdots + a_n v_n : a_i \in [0,1], i = 1,2, \ldots, n\}$$

The volume of this box is defined to be the absolute value of the determinant of the matrix whose rows are the vectors $v_1, v_2, \dots, v_n$.

. . .

The following theorem states some useful properties of determinants.

Theorem 53. Given $n \times n$ matrices A and B, the following properties hold true

(1) $\boldsymbol{det(kA) = k^n\, det(A)}$, $\boldsymbol{k} \in \mathbb{R}$

(2) Assuming A is nonsingular, $\boldsymbol{det(A^{-1}) = \frac{1}{det(A)}}$

(3) $\boldsymbol{det(AB) = det(A)\, det(B)}$

(4) $\boldsymbol{det(A^T) = det(A)}$ ***where*** $\boldsymbol{A^T}$ ***is the transpose of*** $\boldsymbol{A}$

Proof: See Section 3.2 "Properties of Determinants" in the online book by Kuttler [18].

2.13 Cartesian Product

The **Cartesian product** of n sets $A_1, A_2, \dots, A_n$ is the collection of all n-tuples defined as follows:

$$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) : a_i \in A_i, i = 1,2, \dots, n\}$$

As an example of a Cartesian product, we consider multilinear maps. A **multilinear map** is a function of several variables that is linear in each variable taken separately. More precisely, a multilinear map is a function whose domain is the Cartesian product of n vector spaces, and whose codomain is a vector space, i.e., $f : V_1 \times V_2 \times \dots \times V_n \to W$, with the following property:

> For each $i = 1,2, \dots, n$, if all of the variables but v_i are held constant, then $f(v_1, \dots, v_i, \dots, v_n)$ is a linear transformation with respect to v_i.

The above property can be divided into two conditions, i.e.,

$$f(v_1, \dots, v_i + x, \dots, v_n) = f(v_1, \dots, v_i, \dots, v_n) + f(v_1, \dots, x, \dots, v_n), x \in V_i$$

$$f(v_1, \dots, kv_i, \dots, v_n) = kf(v_1, \dots, v_i, \dots, v_n), k \in \mathbb{R}$$

If we view the rows of an $n \times n$ matrix A as vectors in $\mathbb{R}^n$, the $\det A$ is a multilinear map. For a proof of this, see Lemma 3.3 in "Linear Algebra/The Permutation Expansion" [20].

We return to the topic of multilinear maps in Section 2.18.

Cartesian products can be applied to any collection of sets (not necessarily vector spaces). Further, the linearity property in our example is not necessary to define a Cartesian product.

2.14 Similarity, Diagonalization and Eigenvectors

Our definition of equivalent matrices did not require A or B to be square. For example, consider the following 2×3 matrices

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

A and B are equivalent since $B = PAQ$ where

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad Q = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The matrices A and B are said to be **similar** if there is a nonsingular P such that $B = PAP^{-1}$. For matrices to be similar, they need to be square and of the same dimension. So, the concept of similar matrices is a special case of the concept of equivalent matrices. However, the converse is not true. The matrix equivalence class for the $n \times n$ identity matrix I_n consists of all nonsingular matrices of size $n \times n$, since if A^{-1} exist for some $n \times n$ matrix A, then A is equivalent to I_n because $I_n = I_n A A^{-1}$ (noting that $P = I_n$ and $Q = A^{-1}$ in the definition of equivalent matrices). On the other hand, the only member of the similarity class of an identity matrix is the identity matrix itself. To see this, assume A is similar to I, i.e., $I = PAP^{-1}$ for some invertible matrix P. This implies $P = PA$ which, in turn, implies $A = I$.

For example, the matrices below are equivalent but not similar.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

Theorem 54. Matrix similarity is an equivalence relationship.

Proof: **Reflexivity**: Every matrix A is similar to itself since $A = I_n A I_n^{-1}$.

Symmetry: Suppose that $B = PAP^{-1}$. Multiplying both sides on the left by P^{-1} and on the right by P, we have

$$P^{-1}BP = P^{-1}(PAP^{-1})P = A$$

Letting $Q = P^{-1}$, the above can be written as $A = QBQ^{-1}$.

Transitivity: Suppose $A = QBQ^{-1}$ and $B = PCP^{-1}$. Substituting B into the first equation and noting that $(QP)^{-1} = P^{-1}Q^{-1}$ gives us

$$A = Q(PCP^{-1})Q^{-1} = (QP)C(QP)^{-1}$$

Thus, A is similar to C. ■

. . .

If $B = PAP^{-1}$, then $B^n = PA^nP^{-1}, n \in \mathbb{N}$. The proof follows easily by mathematical induction.

. . .

A **diagonal matrix** is a square matrix in which the entries outside the main diagonal are all zero, e.g.,

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}$$

A matrix is said to be **diagonalizable** if it is similar to a diagonal matrix, i.e., A is diagonalizable if there is a nonsingular P such that PAP^{-1} is a diagonal matrix D. A linear transformation $f: V \to V$ is diagonalizable if it has a diagonal representation with respect to the same basis for its domain and codomain.

Example 1. The matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$ is diagonalizable since

$$PAP^{-1} = \begin{bmatrix} 0 & .5 & .5 \\ 1 & 0 & -.5 \\ 0 & -.5 & .5 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} .5 & 1 & .5 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = D$$

If an $n \times n$ matrix A is diagonalizable, i.e., $D = PAP^{-1}$, then we can write $A = P^{-1}DP$. This greatly simplifies the computation of raising A to a power, since $A^k = P^{-1}D^kP$ and

$$D^k = \begin{bmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & d_n^k \end{bmatrix} \text{ if } D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & d_n \end{bmatrix}$$

The following theorem gives us a condition under which a linear transformation is diagonalizable.

Theorem 55. A linear transformation $f: V \to V$ and its representation matrix are diagonalizable if and only if there is a basis $S = \{s_1, s_2, \dots, s_n\}$ and scalars $\lambda_1, \lambda_2, \dots, \lambda_n$ such that $f(s_i) = \lambda_i s_i, i = 1, 2, \dots, n$.

Proof: If a basis as described in the statement of the theorem exists, then we have

$$Rep_{S,S}(f) = \left[Rep_S\big(f(s_1)\big)\ Rep_S\big(f(s_2)\big) \dots Rep_S\big(f(s_n)\big)\right] = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

So, by definition, f is diagonalizable.

Going in the other direction, if f is diagonalizable then there exists a basis S such that

$$Rep_{S,S}(f) = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

This implies that $Rep_S\big(f(s_i)\big) = \begin{bmatrix} 0 \\ \dots \\ \lambda_i \\ \dots \\ 0 \end{bmatrix}_S \Rightarrow f(s_i) = \lambda_i s_i$. ■

. . .

A linear transformation $f: V \to V$ has a scalar **eigenvalue** λ if there is a nonzero **eigenvector** $\vec{\zeta} \in V$ such that $f(\vec{\zeta}) = \lambda\vec{\zeta}$. A square matrix A has a scalar eigenvalue λ associated with the nonzero eigenvector $\vec{\zeta}$ if $A\vec{\zeta} = \lambda\,\vec{\zeta}$.

Example 2. Consider the linear transformation represented by the matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$. The eigenvalues are $1,2,3$ with respective eigenvectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} .5 \\ 1 \\ 1 \end{bmatrix}$. (We are only giving the answer here and will get to the method of solution in the following discussion.)

As a check, note that

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} .5 \\ 1 \\ 1 \end{bmatrix} = 3\begin{bmatrix} .5 \\ 1 \\ 1 \end{bmatrix}$$

. . .

The **characteristic polynomial** of a square matrix A is $\det(A - \lambda I)$ where λ is a variable. The **characteristic equation** is $\det(A - \lambda I) = 0$. The characteristic polynomial of a linear transformation $f: V \to V$ is the characteristic polynomial of any matrix representation $Rep_{S,S}(f)$.

Example 2 (continued). Rewrite the equation $A\vec{\zeta} = \lambda\vec{\zeta}$ as $A\vec{\zeta} - \lambda\vec{\zeta} = (A - \lambda I)\vec{\zeta} = \vec{0}$. This equation has a non-zero solution iff $A - \lambda I$ is singular (i.e. does not have an inverse) iff $\det(A - \lambda I) = 0$. So, we need to solve the characteristic equation for λ.

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

So, the eigenvalues are $1,2,3$.

To find the eigenvector associated with $\lambda = 1$, we plug $\lambda = 1$ into $(A - \lambda I)\vec{\zeta} = \vec{0}$ and solve for $\vec{\zeta}$, i.e.,

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This gives us the equations

$$\zeta_2 = 0$$

$$\zeta_2 + \zeta_3 = 0 \Rightarrow \zeta_3 = 0$$

ζ_1 is a free variable

So, $\zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} = \begin{bmatrix} \zeta_1 \\ 0 \\ 0 \end{bmatrix} = \zeta_1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, i.e., the eigenvector corresponding to $\lambda = 1$ is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

The eigenvectors corresponding to $\lambda = 2$ and $\lambda = 3$ can be found in a similar manner.

Theorem 56. Similar matrices have the same characteristic polynomial and thus, the same eigenvalues.

Proof: Assume that $n \times n$ matrix A is similar to matrix B, and so, there exists some matrix P such that $A = PBP^{-1}$. We have that

$$A - \lambda I_n = PBP^{-1} - \lambda PP^{-1} = PBP^{-1} - P\lambda P^{-1}$$

$$= PBP^{-1} - P\lambda I_n P^{-1} = P(B - \lambda I_n)P^{-1}$$

Thus, $\det(A - \lambda I_n) = \det(P(B - \lambda I_n)P^{-1}) = \det(P)\det(B - \lambda I_n)\det(P^{-1}) = \det(B - \lambda I_n)$.

In the above, we used the fact that $\det P = \frac{1}{\det(P^{-1})}$ (from part (2) of Theorem 53). We also used part (3) of Theorem 53. ■

The converse of Theorem 56 is not true. For example, the matrices $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ have the same characteristic polynomial $(\lambda - 1)^2$, but they are not similar, since the only matrix similar to I_2 is itself.

Theorem 57. If A is similar to B such that $A = PBP^{-1}$ and v is an eigenvector of A, then $P^{-1}v$ is an eigenvector of B corresponding to the same eigenvalue. Further, if u is an eigenvector of B, then Pu is an eigenvector of A corresponding to the same eigenvalue.

Proof: If v is an eigenvector of A such that $Av = \lambda v$, then we have

$$B(P^{-1}v) = (P^{-1}P)B(P^{-1}v) = P^{-1}(PBP^{-1}v) = P^{-1}(Av) = P^{-1}\lambda v = \lambda(P^{-1}v)$$

Thus, $P^{-1}v$ is an eigenvector of B with eigenvalue λ.

If u is an eigenvector of B such that $Bu = \delta u$, then we have

$$A(Pu) = (PBP^{-1})Pu = PBu = P(\delta u) = \delta(Pu)$$

Thus, Pu is an eigenvector of A with eigenvalue δ. ■

The **eigenspace** of a linear transformation $f: V \to V$ associated with the eigenvalue λ is the set

$$\mathcal{E}_\lambda = \{\zeta : f(\zeta) = \lambda\zeta\}$$

The eigenspace of a matrix is defined in an analogous manner. The eigenspace includes the zero vector.

Theorem 58. The eigenspace $\mathcal{E}_\lambda$ of a linear transformation $f: V \to V$ with respect to eigenvalue λ is a subspace of V.

Proof: Since $f(\vec{0}) = \vec{0} = \lambda\,\vec{0},\ \vec{0} \in \mathcal{E}_\lambda$. Since $\mathcal{E}_\lambda$ is not empty, we can use Theorem 2 to show that $\mathcal{E}_\lambda$ is a subspace of V.

Assume $u, v \in \mathcal{E}_\lambda$ and take any linear combination of u and v, $au + bv$. We have

$$f(au + bv) = af(u) + bf(v) = a\lambda u + b\lambda v = \lambda(au + bv)$$

Thus, $au + bv \in \mathcal{E}_\lambda$. ■

Example 3. From Example 1 of this section, we know that $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ are similar matrices. By inspection, we see that the eigenvalues of D are $-1,1,3$. To find the eigenvector associated with the eigenvalue 3, we need to solve

$$(D - 3I)v = 0, \qquad v^T = (x, y, z)$$

$$\Rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow y = z = 0 \text{ and } x \text{ is a free variable.}$$

So, $v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is the eigenvector of D corresponding to eigenvalue 3.

In the equation $PAP^{-1} = D$, we previously determined (in Example 1) that $P^{-1} = \begin{bmatrix} .5 & 1 & .5 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{bmatrix}$.

By Theorem 57, $P^{-1}v = \begin{bmatrix} .5 & 1 & .5 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} .5 \\ 1 \\ 1 \end{bmatrix}$ is the eigenvector of A corresponding to the eigenvalue 3. Let's verify that

$$(A - 3I)(P^{-1}v) = \begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 2 \\ 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} .5 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0}$$

$$\Rightarrow A(P^{-1}v) = 3(P^{-1}v)$$

. . .

If the characteristic polynomial of a matrix factors as $(x - \lambda_1)^{m_1}(x - \lambda_2)^{m_2} \ldots (x - \lambda_k)^{m_k}$ then the eigenvalue λ_i is said to have **algebraic multiplicity** m_i. Its **geometric multiplicity** is the dimension of the associated eigenspace $\mathcal{E}_{\lambda_i}$.

The geometric multiplicity cannot exceed the algebraic multiplicity for a given eigenvalue.

For example, $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ has eigenvalue 2 of algebraic multiplicity 2 and associated eigenspace of dimension 2, i.e., the geometric multiplicity is also 2.

For an $n \times n$ matrix A, $\det A$ is the product of its eigenvalues (counted to multiplicity). [19]

. . .

Theorem 59. For any set of distinct eigenvalues of a linear transformation $f: V \to V$ (or associated $n \times n$ matrix), a set of associated eigenvectors (limited to one per eigenvalue) is linearly independent.

Proof: The plan is to use mathematical induction on the number of eigenvalues, assuming a fixed value for the dimension n of V. The initial step is $k = 0$, i.e., no eigenvalues. In that case, the set of associated vectors is empty and is therefore linearly independent (by definition).

For the inductive step assume that the statement is true for any set of $k > 0$ distinct eigenvalues. Assume f has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{k+1}$ and associated eigenvectors $v_1, v_2, \ldots, v_{k+1}$. Suppose that

$$a_1 v_1 + \cdots + a_k v_k + a_{k+1} v_{k+1} = \vec{0} \quad (1)$$

Multiple Equation (1) by λ_{k+1} to get

$$a_1 \lambda_{k+1} v_1 + \cdots + a_k \lambda_{k+1} v_k + a_{k+1} \lambda_{k+1} v_{k+1} = \vec{0} \quad (2)$$

Apply f to Equation (1) to get

$$a_1 f(v_1) + \cdots + a_k f(v_k) + a_{k+1} f(v_{k+1}) = a_1 \lambda_1 v_1 + \cdots + a_k \lambda_k v_k + a_{k+1} \lambda_{k+1} v_{k+1} = \vec{0} \quad (3)$$

Subtracting Equation (3) from (2), we get

$$a_1(\lambda_{k+1} - \lambda_1)v_1 + \cdots + a_k(\lambda_{k+1} - \lambda_k)v_k + a_{k+1}(\lambda_{k+1} - \lambda_{k+1})v_{k+1} = \vec{0}$$

The v_{k+1} has coefficient 0. Further, from the induction hypothesis, the other coefficients are also 0, i.e., $a_i(\lambda_{k+1} - \lambda_i) = 0, i = 1,2, \ldots, k$. By assumption, the eigenvalues are distinct. So, it must be that $a_i = 0, i = 1,2, \ldots, k$ and so, Equation (1) reduces to $a_{k+1} v_{k+1} = \vec{0}$ which implies $a_{k+1} = 0$. Thus, we have shown that the eigenvectors associated with $k + 1$ eigenvalues are linearly independent, and so, the induction proof is complete. ■

Theorem 60. An $n \times n$ matrix A with n distinct eigenvalues is diagonalizable.

Proof: The n eigenvectors associated with the n eigenvalues are linearly independent by Theorem 59, and thus, form a basis for the column space of A. (By definition, each eigenvalue has at least one nonzero eigenvalue.) By Theorem 55, A is diagonalizable. ■

Theorem 61. An $n \times n$ diagonalizable matrix A has n linearly independent eigenvectors.

Proof: If A is diagonalizable, then there exists an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1} \Rightarrow AP = PD$. Name the columns of P as $v_1, v_2, \ldots, v_n$. We have that $Av_i = \lambda_i v_i, i = 1,2, \ldots, n$ where λ_i is the entry (i, i) in D. So, each v_i is an eigenvector of A corresponding to λ_i. Since P is invertible, its columns $v_1, v_2, \ldots, v_n$ are linearly independent. ■

2.15 Nilpotence

A **nilpotent matrix** is a square matrix M such that $M^k = 0$ for some $k \in \mathbb{N}$. The smallest such positive integer is referred to as the index of M. The following matrix is nilpotent with index 4.

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In general, any upper or lower triangular matrix with zeros on the main diagonal is nilpotent. For additional examples of nilpotent matrices, see the Wikipedia article “Nilpotent matrix”. [21]

If v is an eigenvector of A, i.e., $Av = \lambda v$, then $A^k v = \lambda^k v$. So, if $A^k = 0$, then $\lambda^k v = 0 \Rightarrow \lambda^k = 0 \Rightarrow \lambda = 0$. Thus, every eigenvalue of a nilpotent matrix is zero.

The rank of a nilpotent matrix A decreases with power, i.e., $rank(A) \geq rank(A^2) \geq rank(A^3) \geq \cdots$ and eventually, $rank(A^j) = 0, j \geq k$ where k is the index of A.

The determinant of a nilpotent matrix A is zero. This follows since $A^k = 0 \Rightarrow \det A^k = 0 \Rightarrow (\det A)^k = 0 \Rightarrow \det A = 0$.

A shift matrix is a binary matrix with ones only on the superdiagonal or subdiagonal, and zeroes elsewhere. The matrices below are $\boldsymbol{n} \times \boldsymbol{n}$ upper shift and lower shift matrices, respectively.

$$\boldsymbol{\mathcal{U}_n} = \begin{bmatrix} \mathbf{0} & \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix}, \quad \boldsymbol{\mathcal{L}_n} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \dots & \mathbf{0} & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{1} & \mathbf{0} \end{bmatrix}$$

As a linear transformation, the upper shift matrix "shifts" the components of a vector one position to the left, with a zero appearing in the last position:

$$\boldsymbol{\mathcal{U}_n}(\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_n})^{\boldsymbol{T}} = (\boldsymbol{x_2}, \dots, \boldsymbol{x_n}, \mathbf{0})^{\boldsymbol{T}}$$

Every nilpotent matrix is similar to a block diagonal matrix consisting of upper shift matrices (possibly of different sizes), i.e.,

$$\begin{bmatrix} \boldsymbol{\mathcal{U}_{n_1}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\mathcal{U}_{n_2}} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{\mathcal{U}_{n_k}} \end{bmatrix}$$

Similar statements can be made for lower shift matrices.

Theorem 62. A nilpotent matrix (other than the zero matrix) is not diagonalizable.

Proof: Using proof by contradiction, assume A is diagonalizable. Thus, there exists an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

The diagonal entries of D are the eigenvalues of A (by Theorem 56), but as we have shown, the only eigenvalue of a nilpotent matrix is 0. Therefore, $D = 0$. This implies that $A = P0P^{-1} = 0$, contradicting the assumption that A is nonzero. ■

. . .

A **nilpotent transformation** is a linear transformation f of a vector space V such that $f^{(k)} = 0$ for some $k \in \mathbb{N}$ and thus, $f^{(j)} = 0$ for all $j \geq k$. The notation $f^{(k)}$ means $f(f(\dots))$, i.e., the composition of f with itself k times.

For example, let $\mathcal{P}_n$ be the vector space of single variable polynomials of degree n or less, and let $f: \mathcal{P}_n \to \mathcal{P}_{n-1}$ be the derivative function. If we apply the derivative $n+1$ times, we get the zero function, i.e. $f^{(n+1)} = 0$, and so, f is a nilpotent transformation with index $n+1$.

Theorem 63. A linear transformation on a nontrivial vector space is nilpotent if and only if its only eigenvalue is zero.

Proof: See Lemma 2.1 in Chapter 5, Section IV of Heffron [22]. ■

2.16 Jordan Canonical Form

A **defective matrix** is a square matrix that does not have a complete basis of eigenvectors, meaning it is not diagonalizable. This occurs when at least one eigenvalue has a geometric multiplicity (number of linearly independent eigenvectors) that is strictly less than its algebraic multiplicity (multiplicity in the characteristic polynomial).

Example 1. The following matrix is defective:

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

The eigenvalues are 1 (of algebraic multiplicity 2) and 2 (of algebraic multiplicity 1).

The eigenvalue at 1 has a single eigenvector, i.e., $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and the eigenvalue at 2 has the eigenvector $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. So, by Theorem 61, matrix A is not diagonalizable. However, it is possible to find a matrix J, known as Jordan canonical form of A, such that $A = PJP^{-1}$ for some invertible matrix P.

. . .

The **Jordan normal form** (also referred to as the Jordan canonical form) of a square matrix is a block diagonal matrix composed of Jordan blocks. It represents a matrix in a nearly diagonalized form in cases where a matrix is not diagonalizable.

A Jordan block is a square matrix of the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{bmatrix}$$

The Jordan normal form has the general structure shown below.

$$J = \begin{bmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & J_k \end{bmatrix}$$

Any square matrix over a field $\mathbb{F}$ (in our case, $\mathbb{R}$) has a Jordan normal form if the field is extended to one containing all the eigenvalues of the matrix (for $\mathbb{R}$, this extension is typically $\mathbb{C}$). For example, the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has characteristic equation $\det(A - \lambda I) = \lambda^2 + 1 = 0$ and so, the eigenvalues are $\lambda = \pm i$. Thus, we need to extend to the complex numbers in this case.

The Jordan normal form is not entirely unique, as it is a block diagonal matrix composed of Jordan blocks whose order is not fixed. However, it is unique up to permutation of the blocks. By convention, blocks for the same eigenvalue are grouped together, but no ordering is imposed among the eigenvalues or among the blocks for a given eigenvalue (though one could, for example, order the blocks by weakly decreasing size).

The number of Jordan blocks for a given eigenvalue is equal to the geometric multiplicity of that eigenvalue. The sum of the sizes of all Jordan blocks for a given eigenvalue equals the algebraic multiplicity of that eigenvalue.

A matrix A is diagonalizable if and only if all its Jordan blocks are of size 1×1. In this case, the Jordan normal form reduces to a diagonal matrix.

The Jordan normal form simplifies computations of matrix functions such as the powers A^k and the matrix exponential e^A, since each Jordan block can be treated separately.

. . .

Example 2. The following matrix is defective.

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

The eigenvalues 1 and 2 are of algebraic multiplicity 1, and the eigenvalue 4 is of algebraic multiplicity 2. The geometric multiplicity of the eigenvalue 4 is only 1 and thus, A is not diagonalizable. However, there is a Jordan normal form J and invertible matrix P such that $J = PAP^{-1}$.

$$J = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \qquad P = \begin{bmatrix} -1 & 3 & 1 & 1 \\ 1 & -4 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 5 & 1 & 0 \end{bmatrix}$$

So, there are 3 Jordan blocks for this example. Two of the Jordan blocks are of dimension 1×1 and the other Jordan block (involving the eigenvalue 4) is of dimension 2×2.

Credits: The computations for this example were done using the online app from dCode at https://www.dcode.fr/matrix-jordan.

. . .

Theorem 64. The determinant of a matrix equals the product of its eigenvalues (counting multiplicities).

Proof: Any square matrix A over an algebraically closed field (e.g., the complex numbers) can be transformed into its Jordan canonical form J (shown below) where P is an invertible matrix and J is a block-diagonal matrix consisting of Jordan blocks:

$$A = PJP^{-1}$$

The determinant of A is equal to the determinant of its Jordan form J, since similar matrices have the same determinant, i.e.,

$$\det(A) = \det(PJP^{-1}) = \det(P)\det(J)\det(P^{-1}) = \det(P)\det(J)\frac{1}{\det(P)} = \det(J)$$

Because J is block-diagonal, its determinant is the product of the determinants of the Jordan blocks $\{J_1, J_2, \dots, J_k\}$, i.e.,

$$det(J) = \prod_{i=1}^{k} \det(J_i)$$

The determinant of each Jordan block J_i is simply $\lambda_i^{m_i}$, where m_i is the size of the block, since the determinant of a triangular matrix is the product of its diagonal entries. Thus, we have

$$\det(J_i) = \lambda_i^{m_i}$$

Therefore,

$$\det(A) = \prod_{i=1}^{k} \det(J_i) = \prod_{i=1}^{k} \lambda_i^{m_i}$$

...

Exercises

1. What does the Jordan normal form look like for a 6x6 matrix with eigenvalue 1 (algebraic multiplicity 1), eigenvalue 7 (algebraic multiplicity 1) and eigenvalue 11 (algebraic multiplicity 4 and geometric multiplicity 1).
2. In the previous example, what if the eigenvalue 11 has algebraic multiplicity 4 but geometric multiplicity 2? **Hint**: There are two possibilities.

2.17 Bilinear Maps

The concept of a **bilinear map** will arise when we discuss the Kronecker product of matrices, and the second derivatives in the matrix calculus section.

Let V, W, X be vector spaces over $\mathbb{R}$. A bilinear map is a function $B: V \times W \to X$ such that the following conditions hold true.

- For each fixed $w \in W$, the map $B_w: V \to X$ defined by $B_w(v) = B(v, w)$ is a linear transformation, i.e., $B_w(av_1 + bv_2, w) = aB_w(v_1, w) + bB_w(v_2, w), a, b \in \mathbb{R}$.
- For each fixed $v \in V$, the map $B_v: W \to X$ defined by $B_v(w) = B(v, w)$ is a linear transformation, i.e., $B_v(v, aw_1 + bw_2) = aB_v(v, w_1) + bB(v, w_2), a, b \in \mathbb{R}$.

Theorem 65. For an n-dimensional vector space V, a bilinear form $B: V \times V \to \mathbb{R}$ can be represented by a matrix A relative to a basis of V, i.e., $\{b_1, b_2, \dots, b_n\}$, where $A_{ij} = B(b_i, b_j)$. In particular, for $u, v \in V$, we have that $B(u, v) = u^T Av$.

Proof: Express $u, v \in V$ in terms of the given basis, i.e.,

$$u = \sum_{i=1}^{n} u_i b_i , \qquad v = \sum_{j=1}^{n} v_j b_j$$

Since B is bilinear, we can write

$$B(u, v) = B\left(\sum_{i=1}^{n} u_i b_i , \sum_{j=1}^{n} v_j b_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} u_i v_j B(b_i, b_j)$$

Since $B(b_i, b_j) = A_{ij}$, we have

$$B(u, v) = \sum_{i=1}^{n} \sum_{j=1}^{n} u_i A_{ij} v_j$$

The above is equivalent to $B(u, v) = u^T Av$. ■

Some key points:

- The matrix A compactly encodes the bilinear form B by storing $B(b_i, b_j)$ in its entries.
- The formula $B(u, v) = u^T Av$ allows us to compute B efficiently using matrix operations.
- While the matrix A depends on the choice of basis, the bilinear form B itself is an intrinsic object (**independent of basis**).
- For a fixed basis, the matrix A is unique. A bilinear form can be represented by different matrices in different bases, but they are related by $A' = P^T AP$ where P is a change of basis matrix.
- If $B(u, v) = B(v, u)$ for all $u, v \in V$ then B is symmetric (definition), and its matrix A is symmetric.

For example, let $V = \mathbb{R}^2$ with the standard basis $\{e_1, e_2\}$. Define a bilinear map $B: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ by

$$B(u, v) = 2u_1 v_1 + 3u_1 v_2 + 4u_2 v_1 + 5u_2 v_2, \qquad u = (u_1, u_2), \qquad v = (v_1, v_2)$$

Compute $B(e_i, e_j),\ i = 1,2,\ j = 1,2$

$$B(e_1, e_1) = 2(1)(1) + 3(1)(0) + 4(0)(1) + 5(0)(0) = 2$$

$$B(e_1, e_2) = 3$$

$$B(e_2, e_1) = 4$$

$$B(e_2, e_2) = 5$$

Compute the matrix A, i.e.,

$$A = \begin{bmatrix} B(e_1, e_1) & B(e_1, e_2) \\ B(e_2, e_1) & B(e_2, e_2) \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$$

Let's check that $B(u,v) = u^T Av$.

$$u^T Av = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$= \begin{bmatrix} 2u_1 + 4u_2 & 3u_1 + 5u_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$= (2u_1 + 4u_2)v_1 + (3u_1 + 5u_2)v_2$$

$$= 2u_1v_1 + 3u_1v_2 + 4u_2v_1 + 5u_2v_2 = B(u,v)$$

2.18 Multilinear Map

We can extend the concept of a bilinear map. A **multilinear map** [26] is a function of several variables that is linear separately in each variable. More precisely, a multilinear map is a function $f: V_1 \times V_2 \times \ldots \times V_n \to W$, where $V_1, V_2, \ldots, V_n, W$ are vector spaces, such that function f is a linear transformation of each V_i when the $V_{j \neq i}$ are held constant. Bilinear maps are a special case of multilinear maps.

Let V_i be of dimension d_i for $i = 1,2, \ldots, n$ and W be of dimension d. Let $\{e_{i1}, e_{i2}, \ldots, e_{id_i}\}$ be a basis for V_i and $\{b_1, b_2, \ldots, b_d\}$ be a basis for W. We can define a collection of scalars $A^k_{j_1 \ldots j_n}$ by

$$f(e_{1j_1}, \ldots, e_{nj_n}) = A^1_{j_1 \ldots j_n} b_1 + \cdots + A^d_{j_1 \ldots j_n} b_d = \sum_{k=1}^{d} A^k_{j_1 \ldots j_n} b_k\,, \qquad 1 \leq j_i \leq d_i$$

The above expression represents every possible combination of basis vectors for $V_1 \times V_2 \times \ldots \times V_n$ and the associated mapping to W in terms of the basis for W. Thus, the scalars $A^k_{j_1 \ldots j_n}$ completely determine the multilinear map f. (**Remark**: The superscript on the scalars $A^k_{j_1 \ldots j_n}$ is an index and not a power.)

If we take $v = (v_1, \ldots, v_n) \in V_1 \times V_2 \times \ldots \times V_n$ such that

$$v_i = \sum_{j=1}^{d_i} v_{ij} e_{ij}\,, \qquad 1 \leq i \leq n$$

then

$$f(v) = f(v_1, \ldots, v_n) = f\left(\sum_{j=1}^{d_1} v_{1j} e_{1j}\,, \ldots, \sum_{j=1}^{d_n} v_{nj} e_{nj} \right)$$

$$\sum_{j_1=1}^{d_1} \ldots \sum_{j_n=1}^{d_n} f(e_{1j_1}, \ldots, e_{nj_n})\, v_{1j_1} \ldots v_{nj_n} = \sum_{j_1=1}^{d_1} \ldots \sum_{j_n=1}^{d_n} \sum_{k=1}^{d} A^k_{j_1 \ldots j_n} v_{1j_1} \ldots v_{nj_n} b_k$$

2.19 Dual Spaces

Let V be a vector space over a field $\mathbb{F}$ (commonly $\mathbb{R}$ or $\mathbb{C}$). The dual space of V, denoted by V^*, is the set of all linear transformations $f: V \to \mathbb{F}$.

The dual space V^* is itself a vector space over $\mathbb{F}$ when equipped with addition and scalar multiplication satisfying the following condition:

$$(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x), \forall f, g \in V^*, \qquad x \in V, \qquad \alpha, \beta \in \mathbb{F}$$

Theorem 66. If $\{s_1, s_2, \dots, s_n\}$ is a basis for V, then there exists a basis $\{\theta_1, \theta_2, \dots, \theta_n\}$ of V^* such that $\theta_i(s_j) = \delta_j^i$ where δ_j^i is the Kronecker delta [27].

Proof: Define the linear transformations $\theta_i(a_1 s_1 + a_2 s_2 + \cdots + a_n s_n) = a_i,\ i = 1,2, \dots, n$ for any choice of coefficients from $\mathbb{F}$ (i.e., the field over which V is defined). This gives us the desired condition $\theta_i(s_j) = \delta_j^i$.

The set $\{\theta_i\},\ i = 1,2, \dots, n$ is a basis for V^* because of the following reasons.

1. The $\theta_i,\ i = 1,2, \dots, n$ are functions which map $u, v \in V$ such as $u = \alpha_1 s_1 + \cdots + \alpha_n s_n$ and $v = \beta_1 s_1 + \cdots + \beta_n s_n$ to scalars $\theta_i(u) = \alpha_i$ and $\theta_i(v) = \beta_i$. Further, $\gamma u + \lambda v = (\gamma\alpha_1 + \lambda\beta_1)s_1 + \cdots + (\gamma\alpha_n + \lambda\beta_n)s_n$ and $\theta_i(\gamma u + \lambda v) = \gamma\alpha_i + \lambda\beta_i = \gamma\theta_i(u) + \lambda\theta_i(v)$ which implies $\theta_i \in V^*,\ i = 1,2, \dots, n$, i.e., each θ_i is a linear transformation over the vector space V and field F.

2. Suppose $\lambda_1\theta_1 + \cdots + \lambda_n\theta_n = \mathbf{0} \in V^*$ (where $\mathbf{0}$ is the zero function). We have that $(\lambda_1\theta_1 + \cdots + \lambda_n\theta_n)(s_i) = \lambda_i s_i = 0,\ i = 1,2, \dots, n$ which implies $\lambda_i = 0,\ i = 1,2, \dots, n$ since $s_i \neq \vec{0}, i = 1,2, \dots, n$. So, $\{\theta_1, \dots, \theta_n\}$ is linearly independent in V^*.

3. Finally, take any $f \in V^*$ and $u = \alpha_1 s_1 + \cdots + \alpha_n s_n \in V$. We have that

$$\begin{aligned} f(u) &= f(\alpha_1 s_1 + \cdots + \alpha_n s_n) \\ &= \alpha_1 f(s_1) + \cdots + \alpha_n f(s_n) \\ &= \theta_1(u) f(s_1) + \cdots + \theta_n(u) f(s_n) \\ &= (f(s_1)\theta_1 + \cdots + f(s_n)\theta_n)(u) \end{aligned}$$

   Thus, $\{\theta_1, \dots, \theta_n\}$ generates all of V^* and thus, is a basis for V^*. ■

It follows from the previous theorem, that if V is finite-dimensional, then the dimension of V equals the dimension of V^*.

It is also true that $(V^*)^*$ (or just V^{**}) is isomorphic to V.

Sketch of Proof: If V is finite dimensional, there is a natural isomorphism, i.e., $\phi: V \to V^{**}$ such that $\phi(v) = v^{**}$ where $v^{**}: V^* \to \mathbb{F}$ is defined by $v^{**}(f) = f(v),\ f \in V^*$.

Since the dimensions of V and V^{**} are the same (by Theorem 66), it is sufficient to show that the above natural map ϕ is injective in order to conclude that it is an isomorphism.

Suppose $\phi(v) = 0$. This means $v^{**} = 0$ which implies $v^{**}(f) = 0$ for every $f \in V^*$. By definition of v^{**}, this implies $f(v) = 0$ for all $f \in V^*$. The only vector $v \in V$ satisfying $f(v) = 0$ for all $f \in V^*$ is

$v = 0$ (this is a standard result in linear algebra; it relies on the fact that if $v \neq 0$, we can construct a linear transformation f such that $f(v) \neq 0$). Thus, the null space of ϕ is $\{0\}$ and by Theorem 37, ϕ is injective. ■

Theorem 67. If $\{v_i\}_{i=1}^n$ is a basis for vector space V, and $\{\phi_i\}_{i=1}^n$ is a basis for V^* (the dual of V) such that $\phi_i(v_j) = \delta_j^i$, then $f \in V^*$ can be expressed as

$$f = \sum_{i=1}^{n} f(v_i) \cdot \phi_i$$

Proof: To see this, we evaluate the righthand side of the above at basis vector v_j, i.e.,

$$\sum_{i=1}^{n} f(v_i) \cdot \phi_i(v_j) = \sum_{i=1}^{n} f(v_i) \cdot \delta_j^i = f(v_j)$$

Since righthand and lefthand sides of the above equation agree on each of the basis vectors, they are equal over all of V. ■

A change of basis for the dual of a vector space is related to a change of basis for the vector space as described in the following theorem.

Theorem 68. Let $\{v_i\}_{i=1}^n$ and $\{w_i\}_{i=1}^n$ be bases for vector space V, and $\{\phi_i\}_{i=1}^n$ and $\{\sigma_i\}_{i=1}^n$ be bases for V^* (the dual of V). Assume the bases are selected such that $\phi_i(v_j) = \delta_j^i$ and $\sigma_i(w_j) = \delta_j^i$ (we know this is possible from Theorem 66). If P is the change of basis matrix from $\{v_i\}_{i=1}^n$ to $\{w_i\}_{i=1}^n$ then $\left(P^{-1}\right)^T$ is the change of basis matrix from $\{\phi_i\}_{i=1}^n$ to $\{\sigma_i\}_{i=1}^n$.

Proof: Let $\mathcal{S} = \{v_i\}_{i=1}^n, \mathcal{T} = \{w_i\}_{i=1}^n, \mathcal{A} = \{\phi_i\}_{i=1}^n$ and $\mathcal{B} = \{\sigma_i\}_{i=1}^n$.

To be clear, the matrix P is defined by

$$w_j = \sum_{i=1}^{n} P_{ij} v_i$$

The above implies that column j of P contains the coefficients of w_j in term of $\mathcal{T}$. Using the terminology that we discussed earlier, $P = Res_{\mathcal{T},\mathcal{S}}(i)$.

We seek the matrix Q such that

$$\sigma_j = \sum_{i=1}^{n} Q_{ij} \phi_i$$

The above implies that column j of Q contains the coefficients of σ_j in terms of $\mathcal{A}$. Using the terminology that we discussed earlier, $Q = Res_{\mathcal{B},\mathcal{A}}(i)$.

First, we express $\sigma_j(w_k)$ in terms of Q and P, i.e.,

$$\sigma_j(w_k) = \sum_{i=1}^{n} Q_{ij} \phi_i \left(\sum_{l=1}^{n} P_{lk} v_l \right)$$

Using the linearity of dual vectors, we get

$$\sigma_j(w_k) = \sum_{i=1}^{n}\sum_{l=1}^{n} Q_{ij}P_{lk}\phi_i(v_l)$$

Since $\phi_i(v_l) = \delta_l^i$, the above simplifies to

$$\sigma_j(w_k) = \sum_{i=1}^{n} Q_{ij}P_{ik}$$

Since $\sigma_j(w_k) = \delta_k^j$, the above equation is equivalent to

$$Q^T P = I \Rightarrow Q^T = P^{-1}$$

Taking the transpose on both sides of the above, we get the final result, i.e., $Q = (P^{-1})^T$. ■

3 Matrices

Any fool can know. The point is to understand.

Albert Einstein

3.1 Overview

We've already discussed various aspects of matrices such as determinants, inverses, eigenvalues and eigenvectors and the Jordan canonical form. In this section, we discuss aspects of matrices that are not typically covered in any detail (or not at all) in most introductory linear algebra courses. These aspects are, however, used in matrix calculus and other advanced applications of matrices.

Unlike the previous section, most of the results in the section are stated without proof. Most of the proofs are straightforward but require a lot of space to write out the generic form of the matrices involved in the property and to then show the desired result.

Unless stated otherwise, the matrices and vectors in this section are real-valued.

3.2 Quadratic Forms

A **quadratic form** $Q(x)$ is a polynomial in several variables all of whose terms are of degree 2. A quadratic form can be expressed compactly in terms of an $n \times n$ matrix A and vector $x^T = (x_1, x_2, \dots, x_n)$ as follows:

$$Q(x) = x^T A x$$

The matrix A must be symmetric (i.e., $A = A^T$) for the quadratic form to represent a scalar-valued function with well-defined properties.

Quadratic forms are widely used in optimization, physics (e.g., representing energy), and statistics (e.g., in covariance matrices).

For example, using the matrix $A = \begin{bmatrix} 1 & 3 & 1 \\ 3 & 2 & -1 \\ 1 & -1 & 3 \end{bmatrix}$ and vector $x^T = (x_1, x_2, x_3)$, we can define the following quadratic form:

$$x^T A x = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 3 & 2 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 + 3x_2 + x_3 & 3x_1 + 2x_2 - x_3 & x_1 - x_2 + 3x_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= (x_1^2 + 3x_1x_2 + x_1x_3) + (3x_1x_2 + 2x_2^2 - x_2x_3) + (x_1x_3 - x_2x_3 + 3x_3^2)$$

Note that each term in the above is of degree two, either as a square of a single variable or as the product of two linear terms such as x_2x_3.

A quadratic form is not a linear transformation, e.g., $Q(ax) = (ax)^T A(ax) = a^2 Q(x), a \in \mathbb{R}$ which violates one of the conditions for a linear transformation.

For a linear transformation such as Ax, the equation $Ax = 0$ does not necessarily imply that the matrix A or the vector x equal 0. For example,

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = [0 \quad 0 \quad 0]$$

However, if $Ax = 0, \forall x$, then A must be zero.

The situation is different for quadratic forms, i.e., $x^T Ax = 0, \forall x$ does not imply that $A = 0$ or $x = 0$. For example, consider the following:

$$[\, x_1 \quad x_2 \,] \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [\, -x_2 \quad x_1 \,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -x_1 x_2 + x_1 x_2 = 0$$

Theorem 69. The following properties hold true:

(1) $\boldsymbol{Ax = 0}$ ***for every*** $\boldsymbol{x \iff A = 0}$***,***

(2) $\boldsymbol{x^T Ax = 0}$ ***for every*** $\boldsymbol{x \iff A}$ ***is skew-symmetric (i.e.,*** $\boldsymbol{A^T = -A}$***)***

(3) $\boldsymbol{x^T Ax = 0}$ ***for every*** $\boldsymbol{x}$ ***and*** $\boldsymbol{A = A^T \iff A = 0}$***.***

Proof: We'll prove these properties going in the forward direction only, since going in the other direction is trivial.

(1) Let $x = e_i$ denote the column vector with one in the position i and zeros elsewhere. We have that $Ae_i = 0$ for every i, but $Ae_i = A_{*i}$ (column i of A), and thus, $A = 0$.

(2) If we let $x = e_i + e_j$, then

$$\begin{aligned} &(e_i + e_j)^T A(e_i + e_j) = 0 \\ &\Rightarrow (e_i^T + e_j^T) A(e_i + e_j) = 0 \\ &\Rightarrow e_i^T Ae_i + e_i^T Ae_j + e_j^T Ae_i + e_j^T Ae_j = 0 \\ &\Rightarrow A_{ii} + A_{ij} + A_{ji} + A_{jj} = 0 \end{aligned}$$

By setting $i = j$ in the above, we get $A_{ii} = A_{jj} = 0$ and so, $A_{ij} + A_{ji} = 0$, i.e., $A^T = -A$.

(3) Since $x^T Ax = 0$, we know by (2) that $A^T = -A$. Further, we are given that $A = A^T$ and so, $A = -A$ which implies $A = 0$. ■

Concerning part (1) of the above theorem, suppose we want to prove that two matrices A and B are equal. The direct approach would be to prove $A_{ij} = B_{ij}$ for every i and j. However, it is sometimes easier to show that $Ax = Bx$ for arbitrary x which, in turn, implies $(A - B)x = 0, \forall x \Rightarrow A - B = 0$.

Part (2) also has useful implications. It implies that if $x^T Ax = x^T Bx, \forall x$, then it is not necessarily true that $A = B$, but it is true that $A + A^T = B + B^T = 0$. Further, if A is symmetric (and B is not

necessarily symmetric) and $x^T Ax = x^T Bx, \forall x$, then it still does not follow that A necessarily equals B, but it does follow that $2A = B + B^T \Rightarrow A = \frac{1}{2}(B + B^T)$.

3.3 Positive Definite Matrices

A symmetric matrix A with real entries is said to be **positive-definite** if the real number $x^T Ax$ is positive for every nonzero real column vector x. More generally, an **Hermitian matrix** B (i.e., a complex matrix equal to its conjugate transpose) is positive-definite if the real number $z^* Bz$ is positive for every nonzero complex column vector z, where z^* denotes the conjugate transpose of z. (Note: Since $(z^* Bz)^* = z^* B^* (z^*)^* = z^* Bz$, and $z^* Bz$ is a scalar, $z^* Bz$ must be real.)

Positive semi-definite matrices are defined similarly, except that $x^T Ax$ (or $z^* Bz$ in the complex case) are required to be positive or zero. Negative-definite and negative semi-definite matrices are defined analogously. A matrix that is not positive semi-definite and not negative semi-definite is sometimes called indefinite. [23]

A matrix A is positive-definite if and only if it satisfies any of the following equivalent conditions.

- A is similar to a diagonal matrix D with positive real entries.
- A is symmetric or Hermitian, and all its eigenvalues are real and positive.
- A is symmetric or Hermitian, and all its leading principal minors [24] are positive.
- There exists an invertible matrix B with conjugate transpose B^* such that $A = B^* B$.

The **conjugate transpose** of a matrix, also known as the Hermitian transpose, of an $m \times n$ complex matrix A is an $n \times m$ matrix obtained by transposing A and applying complex conjugation to each entry.

3.4 Trace

The trace of a square matrix A is the sum of its diagonal elements, i.e.,

$$tr(A) = \sum_{i=1}^{n} A_{ii}$$

The trace only applies to square matrices.

The trace is a linear transformation, i.e., $tr(aA + bB) = a\, tr(A) + b\, tr(B), a, b \in \mathbb{R}$.

The following basic properties hold true:

- $tr(A) = tr(A^T)$
- If A and B are two $m \times n$ matrices, then $tr(A^T B) = tr(AB^T) = tr(B^T A) = tr(BA^T)$
- If A and B are $m \times n$ and $n \times m$ real or complex matrices (whose product is thus a square matrix), respectively, then $tr(AB) = tr(BA)$.
- For column vectors $u, v \in \mathbb{R}^n$, the trace of the outer product (see Section 3.6 below) is equivalent to the inner product, i.e., $tr(v\, u^T) = u^T v$.

The similarity-invariance of the trace holds true, i.e., $tr(A) = tr(P^{-1}AP)$ for any square matrix A and any invertible matrix P of the same dimensions. This is proved as follows:

$$tr(P^{-1}AP) = tr\big(P^{-1}(AP)\big) = tr\big((AP)P^{-1}\big) = tr(A)$$

The trace is invariant under circular shifts, e.g.,

$$tr(ABCD) = tr(DABC) = tr(CDAB) = tr(BCDA)$$

Trace invariance holds true for any product of square matrices. However, arbitrary permutations are not necessarily valid, e.g., it is not true in general that $tr(ADBC) = tr(ABCD)$.

The trace is the sum of the eigenvalues of a square matrix A added to multiplicity, i.e., if an eigenvalue is of algebraic multiplicity k it needs to be added k times. This result holds true even if A is a real matrix and some (or all) of the eigenvalues are complex numbers. This property is a consequence of the existence of the Jordan canonical form, together with the similarity-invariance of the trace discussed above.

3.5 Kronecker Product

If A is an $m \times n$ matrix and B is a $p \times q$ matrix, then the Kronecker product of A with B is the $pm \times qn$ block matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \dots & \dots & \dots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix}$$

For example, take the following matrices

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

The Kronecker product of A with B (and B with A) are given by the following (the bold is to aid in readability). In general, $A \otimes B \neq B \otimes A$.

$$A \otimes B = \begin{bmatrix} \boldsymbol{a} & \boldsymbol{2a} & b & 2b \\ \boldsymbol{3a} & \boldsymbol{4a} & 3b & 4b \\ c & 2c & \boldsymbol{d} & \boldsymbol{2d} \\ 3c & 4c & \boldsymbol{3d} & \boldsymbol{4d} \end{bmatrix}, \; B \otimes A = \begin{bmatrix} \boldsymbol{a} & \boldsymbol{b} & 2a & 2b \\ \boldsymbol{c} & \boldsymbol{d} & 2c & 2d \\ 3a & 3b & \boldsymbol{4a} & \boldsymbol{4b} \\ 3c & 3d & \boldsymbol{4c} & \boldsymbol{4d} \end{bmatrix}$$

3.5.1 Bilinear Map

The Kronecker product is a bilinear map, i.e.,

$$A \otimes (B + C) = (A \otimes B) + (A \otimes C)$$

$$(A + B) \otimes C = (A \otimes C) + (B \otimes C)$$

$$(kA) \otimes B = A \otimes (kB) = k(A \otimes B), k \in \mathbb{R}$$

3.5.2 Associative and Non-commutative Properties

In general, we have that $(A \otimes B) \otimes C = A \otimes (B \otimes C)$.

It is also true that $A \otimes 0 = 0 \otimes A = 0$ but not that $A \otimes B$ equals $B \otimes A$. However, there always exists permutation matrices P and Q such that $B \otimes A = P(A \otimes B)Q$.

A **permutation matrix** is a square binary matrix that has exactly one entry of 1 in each row and each column with all other entries 0. Pre-multiplying an n-row matrix M by a permutation matrix P, forming PM, results in permuting the rows of M, while post-multiplying an n-column matrix M, forming MP, permutes the columns of M. For example,

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

$$PM = \begin{bmatrix} m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \\ m_{11} & m_{12} & m_{13} \end{bmatrix}, \qquad MP = \begin{bmatrix} m_{13} & m_{11} & m_{12} \\ m_{23} & m_{21} & m_{22} \\ m_{33} & m_{31} & m_{32} \end{bmatrix}$$

3.5.3 Mixed-product Property

The mixed-product property holds true for the Kronecker product, i.e.,

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

If A is an $m \times n$ matrix and B is a $p \times q$ matrix, then the following is a direct consequence of the mixed product property

$$A \otimes B = \big(A \otimes I_p\big)(I_n \otimes B) = (I_m \otimes B)\big(A \otimes I_q\big)$$

3.5.4 Inverse

The Kronecker product of matrices A and B is invertible if and only if both A and B are invertible, in which case, we have $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

A similar property also holds for the Moore–Penrose pseudoinverse (see Section 3.12 below), i.e., $(A \otimes B)^{+} = A^{+} \otimes B^{+}$.

3.5.5 Transpose, Eigenvalues, Determinant and Trace

Transposition and conjugate transposition (in the case of matrices with complex elements) are distributed over the Kronecker product, i.e.,

$$(A \otimes B)^{T} = A^{T} \otimes B^{T}$$

$$(A \otimes B)^{*} = A^{*} \otimes B^{*}$$

$$\ldots$$

Suppose A is an $n \times n$ matrix and B is an $m \times m$ matrix. If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A and $\mu_1, \ldots, \mu_m$ are the eigenvalues of B then the eigenvalues of $(A \otimes B)$ are

$$\lambda_i \mu_j, \qquad i = 1, \ldots, n, \qquad j = 1, \ldots, m$$

It then follows that the trace and determinant of the Kronecker product of A with B are given by

$$tr(A \otimes B) = tr(A)tr(B)$$

$$\det(A \otimes B) = (\det A)^m (\det B)^n$$

The second result above also relies on the fact that the determinant of a matrix is the product of its eigenvalues (counted to multiplicity). [19]

3.5.6 Kronecker sum and exponentiation

A is an $n \times n$ matrix and B is an $m \times m$ matrix, the Kronecker sum of A and B is defined as

$$A \oplus B = A \otimes I_m + I_n \otimes B$$

We have the following formula concerning the exponential of the Kronecker sum of A and B

$$\exp(A \oplus B) = \exp(\mathrm{A}) \otimes \exp(\mathrm{B})$$

3.6 Outer Product

The **outer product** takes two vectors as input and outputs a matrix . The entries of this matrix are formed by multiplying each element of the first vector by each element of the second vector, resulting in an $n \times m$ matrix if the first vector has n elements and the second has m elements. There is nothing new here beyond the definition of matrix multiplication other than the new term "outer product" for multiplying an $n \times 1$ matrix (column vector) times a $1 \times m$ matrix (row vector).

$$\begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \dots & u_m \end{bmatrix} = \begin{bmatrix} v_1u_1 & v_1u_2 & \dots & v_1u_m \\ v_2u_1 & v_2u_2 & \dots & v_2u_m \\ \dots & \dots & \dots & \dots \\ v_nu_1 & v_nu_2 & \dots & v_nu_m \end{bmatrix}$$

3.7 Variations on Matrix Multiplication

3.7.1 Inner Product Method

The usual matrix multiplication method taught in introductory linear algebra class is known as the inner product approach.

Let A and B be $m \times n$ and $n \times p$ matrices, respectively. Let A_{i*} represent row i of A, and B_{*j} represent column j of B. The product AB is defined to be the $m \times p$ matric C whose (i, j) entry is given by

$$c_{ij} = A_{i*}B_{*j} = \sum_{k=1}^{n} A_{ik}B_{kj}$$

The above is the inner product of row i of A with column j of B.

3.7.2 Linear Combination of Columns Method

Define A and B as above. In this method, column j of the product $C = AB$ is given by

$$C_{*j} = B_{1j}A_{*1} + B_{2j}A_{*2} + \cdots + B_{nj}A_{*n}$$

In words, the j^{th} column of C is the linear combination of the columns of A, with the coefficients being the elements in the j^{th} column of B.

For example, consider the following matrices

$$A = \begin{bmatrix} 1 & 5 & 7 \\ 2 & 1 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 1 & -1 \\ 3 & 3 & 2 \\ 5 & 1 & 4 \end{bmatrix}$$

Using the linear combination of columns method, the first column of $C = AB$ is

$$C_{*1} = B_{11}A_{*1} + B_{21}A_{*2} + B_{31}A_{*3} = 2\begin{bmatrix}1\\2\end{bmatrix} + 3\begin{bmatrix}5\\1\end{bmatrix} + 5\begin{bmatrix}7\\3\end{bmatrix} = \begin{bmatrix}52\\22\end{bmatrix}$$

This and the following two options are not used so much as a computation method but rather as a format in some proofs concerning matrices.

3.7.3 Outer Product Method for Matrices

Define A and B as above. In this method, the product $C = AB$ is given by

$$C = A_{*1}B_{1*} + A_{*2}B_{2*} + \cdots + A_{*n}B_{n*}$$

In words, C is the sum of all the outer product matrices from multiplying each column of A times the corresponding row of B.

Taking A and B from the example in the previous section and applying the outer product method, we have

$$AB = \begin{bmatrix}1\\2\end{bmatrix}\begin{bmatrix}2 & 1 & -1\end{bmatrix} + \begin{bmatrix}5\\1\end{bmatrix}\begin{bmatrix}3 & 3 & 2\end{bmatrix} + \begin{bmatrix}7\\3\end{bmatrix}\begin{bmatrix}5 & 1 & 4\end{bmatrix}$$

$$= \begin{bmatrix}2 & 1 & -1\\4 & 2 & -2\end{bmatrix} + \begin{bmatrix}15 & 15 & 10\\3 & 3 & 2\end{bmatrix} + \begin{bmatrix}35 & 7 & 28\\15 & 3 & 12\end{bmatrix} = \begin{bmatrix}52 & 23 & 37\\22 & 8 & 12\end{bmatrix}$$

3.7.4 Linear Combination of Rows Method

Define A and B as above. In this method, the row i of $C = AB$ is given by

$$C_{i*} = A_{i1}B_{1*} + A_{i2}B_{2*} + \cdots + A_{in}B_{n*}$$

In words, row i of C is gotten by taking the linear combination of all the rows of B, with the coefficients being the elements of row i of A.

Again, using our example from the previous sections, we have

$$C_{1*} = 1 \cdot \begin{bmatrix}2 & 1 & -1\end{bmatrix} + 5 \cdot \begin{bmatrix}3 & 3 & 2\end{bmatrix} + 7 \cdot \begin{bmatrix}5 & 1 & 4\end{bmatrix} = \begin{bmatrix}52 & 23 & 37\end{bmatrix}$$

$$C_{2*} = 2 \cdot \begin{bmatrix}2 & 1 & -1\end{bmatrix} + 1 \cdot \begin{bmatrix}3 & 3 & 2\end{bmatrix} + 3 \cdot \begin{bmatrix}5 & 1 & 4\end{bmatrix} = \begin{bmatrix}22 & 8 & 12\end{bmatrix}$$

3.8 Hadamard Product

The **Hadamard product** (also known as the element-wise product or Schur product) is a binary operation that takes in two matrices of the same dimensions and returns a matrix of the multiplied corresponding elements. More precisely,

$$(A \odot B)_{ij} = A_{ij}B_{ij}$$

For example,

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

3.9 Cofactor Matrix

A **minor of a matrix** *A* is the determinant of some smaller square matrix generated from A by removing one or more of its rows and columns. Minors obtained by removing just one row and one column from square matrices are known as first minors. The (i,j) cofactor of matrix A (denoted $C_{i,j}$) is the first minor of A formed by removing row i and column j of A (denoted $M_{i,j}$) and then multiplying the first minor by $(-1)^{i+j}$, i.e., $C_{i,j} = (-1)^{i+j} M_{i,j}$.

Given $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 3 & 4 & 5 \end{bmatrix}$, $C_{2,2} = (-1)^{2+2} \det \begin{vmatrix} 1 & 3 \\ 3 & 5 \end{vmatrix} = (1 \cdot 5 - 3 \cdot 3) = -4$

The matrix formed by all of the cofactors of a square matrix A is known as the **cofactor matrix** of A.

For our example, the cofactor of A is

$$\begin{bmatrix} -11 & 2 & 5 \\ 2 & -4 & 2 \\ 5 & 2 & -3 \end{bmatrix}$$

The determinate of a matrix A whose element in the position (i,j) is a_{ij} can be expressed along column j as

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

It is also possible to do the expansion along row i, i.e.,

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

3.10 Adjugate matrix

The transpose of the cofactor matrix for a matrix A is called the **adjugate** matrix and is denoted as $adj(A)$.

It can be proved [25] that the inverse of a matrix A (assuming the inverse exists) is the reciprocal of the determinant times the adjugate matrix, i.e.,

$$A^{-1} = \frac{1}{\det(A)} adj(A)$$

Additional properties of the adjugate can be found at the Wikipedia article “Adjugate matrix” [28].

3.11 Vec Operator

The vec operation takes a matrix as input, and outputs a column vector consisting of the columns of the matrix (going from the first column to the last). If a_{ij} is the element on row i and column j of an $n \times m$ matrix A, then

$$vec(A) = (a_{11}, a_{21}, \ldots, a_{n1},\ \ a_{12}, a_{22}, \ldots, a_{n2}, \ldots, a_{1m}, a_{2m}, \ldots, a_{nm})^T$$

(Note $vec(A)$ is written as the transpose of a row vector above, just to save space.)

Consider the following example:

$$A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}, \qquad vec(A) = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{bmatrix}$$

If A is a symmetric matrix, i.e., $A = A^T$, then clearly, $vec(A) = vec(A^T)$.

Theorem 70. If A and B are matrices of the same dimensions (say $m \times n$) then $vec(A)^T vec(B) = tr(A^T B)$.

Proof: The operation $vec(A)^T vec(B)$ sums the product of corresponding elements in A and B, i.e.,

$$vec(A)^T vec(B) = \sum_j \sum_i a_{ij} b_{ij}$$

On the other hand, we have

$$tr(A^T B) = \sum_j (A^T B)_{jj} = \sum_j \sum_i a_{ij} b_{ij}$$

So, the two expressions are equal. ■

Theorem 71. For (column) vectors u and v, $vec(uv^T) = v \otimes u$.

Proof: Let $v = (v_1, v_2, \dots, v_n)^T$ and $u = (u_1, u_2, \dots, u_m)^T$.

We have that uv^T is an $m \times n$ matrix, with column i equal to $v_i u$.

On the other hand, $v \otimes u$ is the $nm \times 1$ matrix given by

$$v \otimes u = (v_1 u, v_2 u, \dots, v_i u, \dots, v_n u)^T$$

Thus, the two expressions are equal. ■

The following theorem is used frequently when taking the derivatives of various matrix expressions.

Theorem 72. For any matrices A, B, C where the product ABC is defined, the following identity holds true

$$vec(ABC) = (C^T \otimes A) vec(B)$$

Proof: The key to this proof is to represent each column of B as the outer product of two vectors. Let b_i be column i of B and e_i be column i of I. Then $b_i e_i^T$ is a matrix whose i^{th} column is b_i and all other entries are 0. Thus, $B = \sum_i b_i e_i^T$.

Consider $vec\left(Ab_i e_i^T C\right) = vec\left((Ab_i)(e_i^T C)\right) = vec\left((Ab_i)(C^T e_i)^T\right)$, noting that Ab_i and $C^T e_i$ are column vectors. Applying Theorem 71, the mixed-product property for the Kronecker product and then Theorem 71 again, we have that

$$vec\left((Ab_i)(C^T e_i)^T\right) = (C^T e_i)\otimes(Ab_i) = (C^T \otimes A)(e_i \otimes b_i) = (C^T \otimes A)vec(b_i e_i^T)$$

So, $vec\left(Ab_i e_i^T C\right) = (C^T \otimes A)vec(b_i e_i^T)$.

Finally, taking the sum over i, we have

$$vec(ABC) = vec\left(A\left(\sum_i b_i e_i^T\right)C\right) = \sum_i Ab_i e_i^T C$$

$$= \sum_i (C^T \otimes A)vec\left(b_i e_i^T\right) = (C^T \otimes A)\sum_i vec(b_i e_i^T)$$

$$= (C^T \otimes A)vec\left(\sum_i b_i e_i^T\right) = (C^T \otimes A)vec(B)$$

In the computation, we made use (several times) of the general property that $vec(\sum_i M_i) = \sum_i vec(M_i)$ for some set of $m \times n$ matrices $\{M_i\}$.

3.12 Moore-Penrose Inverse

The **Moore–Penrose inverse** (aka pseudoinverse or generalized inverse) of a matrix A (denoted A^+) is the most widely known generalization of the inverse matrix. The pseudoinverse applies even to matrices that are not square or not invertible. Given a rectangular matrix with real or complex entries, its pseudoinverse is unique. It was independently described by E. H. Moore in 1920, Arne Bjerhammar in 1951, and Roger Penrose in 1955.

A common use of the pseudoinverse is to compute a "best fit" (least squares) approximate solution to a system of linear equations that lacks an exact solution. Another use is to find the minimum (Euclidean) norm solution to a system of linear equations with multiple solutions. The pseudoinverse facilitates the statement and proof of results in linear algebra.

For a real (or complex) matrix A, the Moore-Penrose inverse A^+ is the unique matrix satisfying the following four conditions (Moore-Penrose conditions):

- $AA^+A = A$
- $A^+AA^+ = A^+$
- $(AA^+)^* = AA^+$ (i.e., AA^+ is Hermitian). A^* is the complex conjugate of A which is just A^T is the entries of A are real.
- $(A^+A)^* = A^+A$ (i.e., A^+A is Hermitian)

Some key properties of the pseudoinverse are as follows:

- If A is invertible, then $A^+ = A^{-1}$.
- If A is of full column rank (i.e., linearly independent columns), then $A^+ = (A^*A)^{-1}A^*$.
- If A has full row rank (i.e., linearly independent rows), then $A^+ = A^*(AA^*)^{-1}$.
- $(A^+)^+ = A$
- The pseudoinversion commutes with transposition, complex conjugation, and taking the conjugate transpose, i.e., $(A^T)^+ = (A^+)^T, (\bar{A})^+ = \overline{A^+}, (A^*)^+ = (A^+)^*$.
- The pseudoinverse of a scalar multiple of A is the reciprocal multiple of A^+, i.e., $(\alpha A)^+ = \frac{1}{\alpha}A^+, \alpha \neq 0$.
- The pseudoinverse provides the least-squares solution to $Ax = b$, minimizing $\|Ax - b\|$ (the Euclidean norm).

3.13 Permutation Matrix

A **permutation matrix** [29] is a square binary matrix that has a single 1 in each row and each column with all other entries being 0. An $n \times n$ permutation matrix can represent a permutation of n elements. Pre-multiplying an n-row matrix M by a permutation matrix P, forming matrix PM, results in permuting the rows of M, while post-multiplying an n-column matrix M, forming matrix MP, permutes the columns of M.

Every permutation matrix P is orthogonal, i.e., its inverse is equal to its transpose: $P^{-1} = P^T$.

In the example below, multiplication on the left by matrix P, permutates rows 1 and 2 of matrix A.

$$PA = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix}$$

3.14 Commutation Matrix

It is not true in general that $vec(A) = vec(A^T)$. For instance, consider A^T from our example in Section 3.11.

$$A^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \qquad vec(A^T) = \begin{bmatrix} 1 \\ 4 \\ 7 \\ 2 \\ 5 \\ 8 \\ 3 \\ 6 \\ 9 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{bmatrix} = vec(A)$$

While $vec(A) \neq vec(A^T)$, there always exists a matrix $K_{m,n}$ (known as the **commutation matrix**) such that $K_{m,n}vec(A) = vec(A^T)$ where A is an $m \times n$ matrix. $K_{m,n}$ contains mn ones and $mn(mn-1)$ zeros. Further, the commutation matrix is of dimensions $mn \times mn$, with exactly one entry of 1 in each row and each column with all other entries 0 (such matrices are known as permutation matrices, as noted in the previous section). For example,

$$K_{3,3} = \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix}$$

When multiplied times a 9×1 vector, $K_{3,3}$ takes the element in the first position to the first position on the output, the second element to the 4th position, the 3rd element to the 7th position and so on. This works not just for example above but for any vectorization of any 3×3 matrix.

$K_{3,3}$ can be viewed as a matrix consisting of 9 blocks of 3×3 matrices, each with a single 1 and zeros everywhere else. There are three diagonal blocks. The first diagonal block has a 1 in position $(1,1)$, the second diagonal block has a 1 in relative position $(2,2)$ and the third diagonal block has a 1 in relative position $(3,3)$. Thus, $tr(K_{3,3}) = 3$. A similar result holds for $tr(K_{n,n})$.

Theorem 73. The following properties hold true for the commutation matrix: (1) $tr(\boldsymbol{K_{n,n}}) = \boldsymbol{n}$ and (2) $\boldsymbol{K_{m,n}^T = K_{m,n}^{-1} = K_{n,m}}$.

Proof: (1) For an $n \times n$ matrix A, its i^{th} diagonal element is placed in the same position in $vec(A)$ as in $vec(A^T)$. Each of the n diagonal blocks of $K_{n,n}$ has single 1 along its diagonal and thus, $tr(K_{n,n}) = n$.

(2) As noted in Section 3.13, all permutation matrices are orthogonal and so, $K_{m,n}^T = K_{m,n}^{-1}$. Multiplying $K_{m,n}vec(A) = vec(A^T)$ on the left by $K_{n,m}$, we get $K_{n,m}K_{m,n}vec(A) = K_{n,m}vec(A^T) = vec(A)$ which implies $K_{m,n}^{-1} = K_{n,m}$. ■

An important property of the commutation matrix is that it enables us to interchange (commute) the two matrices of a Kronecker product.

Theorem 74. For any $\boldsymbol{m \times n}$ matrix $\boldsymbol{A}$ and $\boldsymbol{p \times q}$ matrix $\boldsymbol{B}$, $\boldsymbol{K_{p,m}(A \otimes B) = (B \otimes A)K_{q,n}}$.

Proof: We prove the theorem by showing that matrices $K_{p,m}(A \otimes B)$ and $(B \otimes A)K_{q,n}$ have the same effect (result) when applied to an arbitrary vector. To that end, let X be an arbitrary $q \times n$ matrix.

By repeated applications of Theorem 72 and the definition of the commutation matrix, we have

$$K_{p,m}(A \otimes B)vec(X) = K_{p,m}vec(BXA^T) = vec(AX^TB^T)$$

$$= (B \otimes A)vec(X^T) = (B \otimes A)K_{q,n}vec(X)$$

Since X was chosen arbitrarily, $K_{p,m}(A \otimes B) = (B \otimes A)K_{q,n}$. ■

Exercise: Show that $N_n = \frac{1}{2}(I + K_{n,n})$ is symmetric idempotent (i.e., $N_n = N_n^T$ and $N_n^2 = I$) and that $N_n(A \otimes A) = (A \otimes A)N_n$ for every $n \times n$ matrix A.

By Theorem 73, $K_{n,n}$ is symmetric and orthogonal, and thus, $K_{n,n}^2 = K_{n,n}K_{n,n}^{-1} = I$. Further,

$$N_n^T = \frac{1}{2}\left(I + K_{n,n}^T\right) = \frac{1}{2}\left(I + K_{n,n}\right) = N_n$$

$$N_n^2 = \frac{1}{4}\left(I + K_{n,n}\right)^2 = \frac{1}{4}\left(I + 2K_{n,n} + K_{n,n}^2\right) = \frac{1}{4}\left(2I + 2K_{n,n}\right) = N_n$$

By Theorem 74, we have

$$K_{n,n}(A \otimes A) = (A \otimes A)K_{n,n}$$

Making use of the above equation,

$$N_n(A \otimes A) = \frac{1}{2}(A \otimes A) + \frac{1}{2}K_{n,n}(A \otimes A) = \frac{1}{2}(A \otimes A) + \frac{1}{2}(A \otimes A)K_{n,n} = (A \otimes A)N_n$$

3.15 Duplication Matrix

Symmetric matrices have redundant information, i.e., the information above the main diagonal is the same as that below the main diagonal. This redundancy carries over to the vectorization of a symmetric matrix. For purposes of computation efficiency, it is desirable to have a non-redundant version of the vec operation for symmetric matrices. This operation is known as $vech$ (short for "vec-half"). For an $n \times n$ matrix A, the $vech$ operation produces a $\frac{1}{2}n(n+1)$ column vector by eliminating all the elements of A above the main diagonal. For example,

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -3 & 5 \\ -1 & 5 & 4 \end{bmatrix}, \qquad vec(A) = \begin{bmatrix} 1 \\ 2 \\ -1 \\ \mathbf{2} \\ -3 \\ 5 \\ \mathbf{-1} \\ \mathbf{5} \\ 4 \end{bmatrix}, \qquad vech(A) = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -3 \\ 5 \\ 4 \end{bmatrix}$$

Since the elements of $vech(A)$ are those of $vec(A)$, without the redundancies, there exists a unique matrix which transforms, for a symmetric $n \times n$ matrix A, $vech(A)$ into $vec(A)$. This matrix is known as the **duplication matrix** and is denoted by D_n. If A is $n \times n$, then D_n is $n^2 \times \frac{1}{2}n(n+1)$. More precisely,

$$D_n vech(A) = vec(A)$$

For example,

$$D_3 = \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix}$$

The duplication matrix D_n has full column rank because its columns are linearly independent by construction.

- Each distinct element a_{ij} of A (where $i \geq j$) is mapped by D_n from $vech(A)$ to one or two positions in $vec(A)$, i.e.,
  - If $i = j$ (diagonal element of A), a_{ii} appears once in $vec(A)$.
  - If $i \neq j$ (off-diagonal element of A), then it appears twice in $vec(A)$, at positions corresponding to a_{ij} and a_{ji}.
- The columns of D_n are structured to only overlap where symmetry necessitates, but they do not redundantly represent the same information.

Consequences of D_n being of full column rank:

- $D_n^T D_n$ is positive definite: For any vector x, we have $x^T(D_n^T D_n)x = (D_n x)^T(D_n x) = \|D_n x\|^2 \geq 0$. This shows $D_n^T D_n$ is positive semi-definite. Since the columns of D_n are linearly independent, $D_n x = 0$ only if $x = 0$. Thus, for any $x \neq 0,\ \|D_n x\|^2 > 0$.
- A positive-definite matrix is always invertible, so $(D_n^T D_n)^{-1}$ always exists.
- The pseudoinverse of D_n is given by $D_n^+ = (D_n^T D_n)^{-1} D_n^T$, which allows us to recover $vech(A)$ from $vec(A)$, i.e.,

$$D_n vech(A) = vec(A)$$
$$D_n^+ D_n vech(A) = D_n^+ vec(A)$$
$$(D_n^T D_n)^{-1}(D_n^T D_n) vech(A) = D_n^+ vec(A)$$
$$vech(A) = D_n^+ vec(A)$$

- In the context of the above equation, D_n^+ is known as the elimination matrix.

Theorem 75. The duplication matrix is related to the commutation matrix by the following identities:

$$\boldsymbol{K_{n,n} D_n = D_n}, \qquad \boldsymbol{D_n (D_n^T D_n)^{-1} D_n^T = D_n D_n^+ = \frac{1}{2}(I + K_{n,n})}$$

Proof: Concerning the first identity, take any symmetric matrix X. Using the definition of D_n and $K_{n,n}$, we have

$$K_{n,n} D_n vech(X) = K_{n,n} vec(X) = vec(X^T) = vec(X) = D_n vech(X)$$

The symmetry of X does not restrict the possible values of $vech(X)$ and so, for arbitrary $vech(X)$, $K_{n,n}D_n vech(X) = D_n vech(X)$ which implies $K_{n,n}D_n = D_n$.

Concerning the second result, let $M_n = D_n(D_n^T D_n)^{-1}D_n^T$, $N_n = \frac{1}{2}(I_{n^2} + K_{n,n})$ and $\Delta_n = N_n - M_n$. In a previous exercise, we showed that N_n is symmetric idempotent. M_n is also symmetric idempotent, as we prove below.

(Symmetric) Applying Theorem 21, we have

$$M_n^T = (D_n(D_n^T D_n)^{-1}D_n^T)^T = D_n((D_n^T D_n)^{-1})^T D_n^T$$
$$= D_n((D_n^T D_n)^T)^{-1}D_n^T = D_n(D_n^T D_n)^{-1}D_n^T = M_n$$

(Idempotent)

$$M_n^2 = [D_n(D_n^T D_n)^{-1}D_n^T][D_n(D_n^T D_n)^{-1}D_n^T] = D_n(D_n^T D_n)^{-1}\boldsymbol{(D_n^T D_n)}(D_n^T D_n)^{-1}D_n^T$$
$$= D_n(D_n^T D_n)^{-1}ID_n^T = D_n(D_n^T D_n)^{-1}D_n^T = M_n^2$$

Using the result that $K_{n,n}D_n = D_n$ (or equivalently, $D_n^T K_{n,n}^T = D_n^T$) and the fact that $K_{n,n}^T = K_{n,n}$ (by Theorem 73), we have that

$$M_n N_n = \frac{1}{2}D_n(D_n^T D_n)^{-1}D_n^T(I + K_{n,n}) = \frac{1}{2}\left[D_n(D_n^T D_n)^{-1}D_n^T + D_n(D_n^T D_n)^{-1}D_n^T K_{n,n}\right]$$

$$= \frac{1}{2}\left[D_n(D_n^T D_n)^{-1}D_n^T + D_n(D_n^T D_n)^{-1}D_n^T K_{n,n}^T\right] = \frac{1}{2}[D_n(D_n^T D_n)^{-1}D_n^T + D_n(D_n^T D_n)^{-1}D_n^T] = M_n$$

Similarly, it can be shown that $N_n M_n = M_n$.

Using the above results, it is easy to verify that Δ_n is also symmetric idempotent.

By Theorem 73, $tr(K_{n,n}) = n$, and so, $tr(N_n) = tr\left(\frac{1}{2}(I_{n^2} + K_{n,n})\right) = \frac{n^2+n}{2}$.

By cyclic property of the trace operator, we have

$$tr(M_n) = tr(D_n(D_n^T D_n)^{-1}D_n^T) = tr(D_n^T D_n(D_n^T D_n)^{-1}) = tr\left(I_{\frac{n(n+1)}{2}}\right) = \frac{n(n+1)}{2}$$

So, $tr(N_n) = tr(M_n)$ which implies that $tr(\Delta_n) = 0$.

In general, the trace of an idempotent matrix is equal to its rank [30] and only the zero matrix has zero rank.

Thus, $tr(\Delta_n) = 0 \Rightarrow \Delta_n = 0 \Rightarrow M_n = N_n$. ■

3.16 Block Matrices

Some matrices have an internal structure (a block structure) that proves useful in deriving various results. We have already seen several examples, i.e., the Jordan canonical form in Section 2.16, the duplication matrix D_n, and the Kronecker product is defined in terms of blocks.

A **block matrix** is a matrix that is partitioned into rectangular submatrices, such that the entire matrix is composed of blocks.

This notion can be made more precise for an $n \times m$ matrix M by partitioning n into a collection of **rowgroups** and partitioning m into a collection of **colgroups**. The matrix M is then considered as the whole of these groups.

In Table 1, we have a block matrix defined by the rowgroups $\{1,2\},\{3,4\},\{5,6,7\},\{8,9,10\}$ and the colgroups $\{1,2,3\},\{4,5\},\{6,7,8\},\{9,10\}$. For example, the rowgroup $\{5,6,7\}$ and colgroup $\{3,4\}$ define a block (highlighted by a bold border in the table).

Table 1. Block matrix defined by rowgroups and colgroups

x	x	x			x	x	x		
x	x	x			x	x	x		
x	x	x			x	x	x		
x	x	x			x	x	x		
x	x	x			x	x	x		
x	x	x			x	x	x		
x	x	x			x	x	x		
x	x	x			x	x	x		
x	x	x			x	x	x		
x	x	x			x	x	x		

Another way to represent the partitioning of a matrix is by explicitly identifying the blocks. As shown below, A is a block matrix consisting of pq block matrices. The blocks in any one column must all have the same number of columns, and the blocks in any one row must have the same number of rows.

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1q} \\ A_{21} & A_{22} & \dots & A_{2q} \\ \dots & \dots & \dots & \dots \\ A_{p1} & A_{p2} & \dots & A_{pq} \end{bmatrix}$$

AWhen represented in the above form and such that all the blocks are of the same dimension, the transpose of A can be represented as the transpose of the blocks, i.e.,

$$A^T = \begin{bmatrix} A_{11}^T & A_{21}^T & \dots & A_{p1}^T \\ A_{12}^T & A_{22}^T & \dots & A_{p2}^T \\ \dots & \dots & \dots & \dots \\ A_{1q}^T & A_{2q}^T & \dots & A_{pq}^T \end{bmatrix}$$

Other types of partitions are possible. For example, we can represent an $p \times q$ matrix A by blocks of columns A_{*j}, i.e., $A = [A_{*1}, A_{*2}, \dots, A_{*q}]$.

Another approach is to represent A by blocks of rows A_{i*}, i.e.,

$$A = \begin{bmatrix} A_{1*} \\ A_{2*} \\ \dots \\ A_{p*} \end{bmatrix}$$

Assume we have a matrix B that is of dimensions $p \times q$ and has the same block structure as matrix A, i.e.,

$$B = \begin{bmatrix} B_{11} & B_{12} & \dots & B_{1q} \\ B_{21} & B_{22} & \dots & B_{2q} \\ \dots & \dots & \dots & \dots \\ B_{p1} & B_{p2} & \dots & B_{pq} \end{bmatrix}$$

In this case, we can describe a linear combination of A and B at the level of their blocks, i.e.,

$$\alpha A + \beta B = \begin{bmatrix} \alpha A_{11} + \beta B_{11} & \alpha A_{12} + \beta B_{12} & \dots & \alpha A_{1q} + \beta B_{1q} \\ \alpha A_{21} + \beta B_{21} & \alpha A_{22} + \beta B_{22} & \dots & \alpha A_{2q} + \beta B_{2q} \\ \dots & \dots & \dots & \dots \\ \alpha A_{p1} + \beta B_{p1} & \alpha A_{p2} + \beta B_{p2} & \dots & \alpha A_{pq} + \beta B_{pq} \end{bmatrix}$$

It's possible to multiple matrices in block form but the blocks need to be compatible for multiplication, i.e.,

- Each block A_{ik} is of size $m_i \times n_k$.
- Each block B_{kj} is of size $n_k \times p_j$.

As defined above, the product of matrix A with matrix B is defined as follows:

$$AB = \begin{bmatrix} \sum_{i=1}^{q} A_{1i}B_{i1} & \sum_{i=1}^{q} A_{1i}B_{i2} & \dots & \sum_{i=1}^{q} A_{1i}B_{iq} \\ \sum_{i=1}^{q} A_{2i}B_{i1} & \sum_{i=1}^{q} A_{2i}B_{i2} & \dots & \sum_{i=1}^{q} A_{2i}B_{iq} \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{q} A_{pi}B_{i1} & \sum_{i=1}^{q} A_{pi}B_{i2} & \dots & \sum_{i=1}^{q} A_{pi}B_{iq} \end{bmatrix}$$

Additional concepts and results concerning block matrices can be found in the Wikipedia article "Block matrix" [31].

3.17 LU Decomposition

Lower–Upper (LU) decomposition factors a matrix into the product of a lower triangular matrix and an upper triangular matrix. The product sometimes includes permutation matrices. LU decomposition can be viewed as the matrix form of Gaussian elimination. Computers usually solve square systems of linear equations using LU decomposition, and it is also a key step when inverting a matrix or computing the determinant of a matrix. The LU decomposition was introduced by the Polish astronomer Tadeusz Banachiewicz in 1938. [32]

Let A be a square matrix. An LU decomposition refers to expression of A into product of two factors – a lower triangular matrix L and an upper triangular matrix U such that $A = LU$. In some case, the decomposition is not possible without prior reordering of A to prevent division by zero or uncontrolled growth of rounding errors. In such cases, the factorization decomposition of A takes the form $PAQ = LU$, where the matrices P and Q perform (via multiplication) permutations on the rows and columns of A. This rearrangement of rows and columns is known as **pivoting**. For a square matrix, it is always possible to perform LU decomposition, but pivoting is required in some cases.

Conditions for the existence and uniqueness of the LU decomposition are stated in the "Existence and uniqueness" section of the Wikipedia article "LU decomposition" [32].

For example, consider the matrix

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

The first step is to replace the 0 in the upper left. There are several options here. We choose to switch rows 1 and 2 which is done via multiplication by the permutation matrix

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Next, we find the L and U matrices by solving the following system of equations:

$$PA = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 2 & 1 \\ 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

Multiplying LU and equating with the entries in PA, we have

$$u_{11} = 1, \quad u_{12} = -2, \quad u_{13} = 0$$

$$l_{21}u_{11} = 0 \Rightarrow l_{21} = 0$$

$$l_{21}u_{12} + u_{22} = 2 \Rightarrow u_{22} = 2$$

$$l_{21}u_{13} + u_{23} = 1 \Rightarrow u_{23} = 1$$

$$l_{31}u_{11} = 2 \Rightarrow l_{31} = 2$$

$$l_{31}u_{12} + l_{32}u_{22} = 3 \Rightarrow 2(-2) + 2l_{32} = 3 \Rightarrow l_{32} = 3.5$$

$$l_{31}u_{13} + l_{32}u_{23} + u_{33} = 4 \Rightarrow 2(0) + (3.5)(1) + u_{33} = 4 \Rightarrow u_{33} = .5$$

Thus, we have

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 3.5 & 1 \end{bmatrix}, U = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & .5 \end{bmatrix}$$

If we started by switching rows 1 and 3, we would have gotten different solutions for L and U.

3.18 QR Decomposition

A **QR decomposition** is a factorization of a matrix A into a product $A = QR$ of an orthonormal matrix Q and an upper triangular matrix R. QR decomposition is often used to solve the linear least squares problem and is the basis for a particular eigenvalue algorithm, the QR algorithm. [33]

Any real square matrix A may be decomposed as $A = QR$, where Q is an orthogonal matrix (its columns are orthogonal unit vectors meaning $Q^T = Q^{-1}$) and R is an upper triangular matrix (also called right triangular matrix). If A is invertible, then the factorization is unique if we require the diagonal elements of R to be positive.

If A is a complex square matrix, then there is a decomposition $A = QR$ where Q is a **unitary matrix**, i.e., $Q^* = Q^{-1}$ where Q^* is the conjugate transpose of Q.

A rectangular matrix A of size $m \times n$ (where $m > n$) always has a QR decomposition as long as its columns are linearly independent, see Theorem 7.2 in "Numerical Linear Algebra" [34] for a proof of this fact. The QR decomposition factorizes $A = QR$ such that

- Q is an $m \times n$ matrix with orthonormal columns (i.e., $Q^T Q = I_n$).
- R is an $n \times n$ upper triangular matrix.

3.19 Singular Value Decomposition

The **singular value decomposition** (SVD) is a factorization of a real or complex matrix into a rotation, followed by a rescaling and then another rotation. [35]

More precisely, given an $m \times n$ matrix A (real or complex), its SVD is a factorization of the form $A = U\Sigma V^*$ where:

- U is an $m \times m$ unitary matrix (orthogonal if A is real) whose columns are the left singular vectors of A.
- Σ (Sigma) is an $m \times n$ rectangular diagonal matrix with non-negative real numbers σ_i (the singular values) on the diagonal, ordered such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, where r is the rank of A.
- V^* (conjugate transpose of V) is an $n \times n$ unitary matrix (orthogonal if A is real) whose columns are the right singular vectors of A.

A non-negative real number σ is a singular value of A if there exist unit vectors $u \in \mathbb{C}^m$ and $v \in \mathbb{C}^n$ such that $Av = \sigma u$ and $A^* u = \sigma v$. The vectors u and v are called left-singular and right-singular vectors for σ, respectively.

Some points of clarification:

1. Regarding the matrix Σ
    a. The diagonal entries of Σ are non-negative and ordered ($\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$).
    b. The remaining entries (if $m \neq n$) are zero. For example, if $m > n$, Σ has n diagonal entries and $m - n$ rows of zeros.
2. Regarding uniqueness
    a. The singular values σ_i are uniquely determined by A.
    b. The singular vectors (i.e., the columns of U and V) are not unique (e.g., they can be multiplied by a unit complex number $e^{i\theta}$ or by -1 in the real case).

3. Reduced versus full SVD
    a. The definition provided above is the "full" SVD version where U and V are square.
    b. There is also a "reduced" SVD, where U is $m \times r, \Sigma$ is $r \times r$, and V is $n \times r$
4. Relation to eigenvalues
    a. The singular values of A are the square roots of the eigenvalues of A^*A (or AA^* which as the same eigenvalues as A^*A).
    b. The left and right singular vectors are eigenvectors of A^*A and AA^*, respectively.

As an example, consider the matrix

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

We use item #4 from the "points of clarification" above to determine the SVD for A.

First, we compute A^TA and AA^T.

$$A^TA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Next, we determine the eigenvalues and eigenvectors of A^TA (to get V and Σ).

$$\det(A^TA - \lambda I) = 0$$

$$\det\begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} = (1-\lambda)[(1-\lambda)(2-\lambda)-1] - 0 + 1[0-(1-\lambda)] = 0$$

$$\lambda(\lambda-1)(\lambda-3) = 0$$

Thus, the eigenvalues of A^TA are $\lambda = 3,1,0$, and the singular values are $\sigma_i = \sqrt{\lambda_i}$ and so,

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Find the eigenvector corresponding to $\lambda_1 = 3$ and normalize the vector (i.e., divide by its length).

$$(A^TA - 3I)v_1 = 0 \Rightarrow \begin{bmatrix} -2 & 0 & 1 \\ 0 & -2 & 1 \\ 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \Rightarrow v_1 = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Find the eigenvector corresponding to $\lambda_2 = 1$ and normalize the vector.

$$(A^TA - I)v_2 = 0 \Rightarrow \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \Rightarrow v_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

Find the eigenvector corresponding to $\lambda_2 = 1$ and normalize the vector.

$$A^T A v_3 = 0 \Rightarrow v_3 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Using the above result, we can determine V, i.e.,

$$V = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \end{bmatrix}$$

The eigenvalues for AA^T are the same as those for $A^T A$.

To determine U, we need to find the normalized eigenvectors of AA^T. We leave the details as an exercise. The result is

$$u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad u_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Thus,

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

We now have all the input to form the SVD for matrix A.

$$A = U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{bmatrix}$$

3.20 Frobenius Norm

The **Frobenius norm** of an $m \times n$ matrix A is defined as the square root of the sum of the absolute squares of its elements, i.e.,

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}$$

It is also equal to the square root of the trace of AA^* (where A^* is the conjugate transpose of A), i.e.,

$$\|A\|_F = \sqrt{tr(AA^*)}$$

Theorem 76. For an $m \times n$ matrix A, the square of the Frobenius norm of A is equal to the sum of the squared Euclidean norms of its columns, i.e.,

$$\|\boldsymbol{A}\|_{\boldsymbol{F}}^{\boldsymbol{2}} = \sum_{\boldsymbol{j=1}}^{\boldsymbol{n}}\|\boldsymbol{a_j}\|^{\boldsymbol{2}}$$

where a_j is column j of A.

Proof: By definition of the Frobenius norm, we have

$$\|\boldsymbol{A}\|_{\boldsymbol{F}}^{\boldsymbol{2}} = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2$$

Switching the order of summation gives us

$$\|\boldsymbol{A}\|_{\boldsymbol{F}}^{\boldsymbol{2}} = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2 = \sum_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|^2 = \sum_{j=1}^{n}\|a_j\|^2$$

4 Matrix Calculus

To not know math is a severe limitation to understanding the world.

Richard P. Feynman

4.1 Scope

The simplest case for derivatives concerns functions with a scalar input and a scalar output. For example, the function $f(x) = x^n, n \in \mathbb{N}$ has derivatives $f'(x) = nx^{n-1}$. Anyone who has taken an introductory calculus course knows this formula.

[For the examples that follow in this subsection, the elements of the vectors and matrices are assumed to be real or complex numbers, unless stated otherwise. Vectors are assumed to be column vectors. If a row vector is required, it will be written as the transpose of a column vector.]

Next, in terms of complexity, we have two branches, i.e., functions with a vector input and scalar output, and functions with a scalar input and vector output. An example of the first case is $f(x, y, z) = x^2 + xy + y^2 + yz + z^2$. An example of the second case is $f(x) = (x^2, 3x, \sqrt{x})$.

From here, we advance to the case of functions that have vector inputs and vector outputs, e.g., $f(x, y, z) = (5xyz, x^2 + z^2)$.

Some additional cases with vector inputs:

- Vector input with scalar output (involving vectors and matrices), e.g.,
  - $f(x) = \|x^T\|$, where the variable x is an $n \times 1$ column vector of real numbers and $\|x^T\|$ is the Euclidean norm of x (i.e., the length of x)
  - $f(x) = x^T x$, where the variable x is an $n \times 1$ column vector
  - $f(x) = x^T A x$, where the variable x is an $n \times 1$ column vector and A is a constant $n \times n$ matrix
- Vector input with vector output (involving vectors and matrices), e.g.,
  - $f(x) = Ax$, where the variable x is an $n \times 1$ column vector and A is a constant $n \times n$ matrix
  - $f(x) = \frac{x}{\|x^T\|}$, where the variable x is an $n \times 1$ column vector of real numbers. This function maps each nonzero vector to a unit vector in the same direction.
- Vector input with matrix output, i.e., $f(x) = xx^T$, where the variable x is an $n \times 1$ column vector and thus, xx^T is an $n \times n$ matrix.

Some cases with matrix inputs:

- Matrix input with scalar output, e.g.,
  - $f(X) = a^T X a$, where a is a constant $n \times 1$ column vector and the variable X is an $n \times n$ matrix.
  - $f(X) = tr(X)$, where the variable X is an $n \times n$ matrix
  - $f(X) = \det(X)$, where the variable X is an $n \times n$ matrix

- Matrix input with vector output, e.g.,
  - $f(X) = Xa$, where a is a constant $n \times 1$ column vector and the variable X is an $n \times n$ matrix.
  - $f(X) = vec(X)$
- Matrix input with matrix output, e.g.,
  - $f(X) = X^n, n \in \mathbb{N}$, where the variable X is an $n \times n$ matrix
  - $f(X) = e^X$, where the variable X is an $n \times n$ matrix

4.2 Derivatives of Scalar Functions

The **gradient** of a scalar-valued differentiable function f with a vector input is the vector-valued function ∇f (stated as "nabla f") whose value at a vector x gives the direction and the rate of fastest increase. If the gradient of a function is non-zero at an input vector (point) x, the direction of the gradient is the direction in which the function increases most quickly from x, and the magnitude of the gradient is the rate of increase in that direction, i.e., the greatest absolute directional derivative. In addition, an input vector (point) where the gradient is the zero vector is known as a **stationary point**.

For functions from $\mathbb{R}^n$ to $\mathbb{R}$, the gradient is given by the set of partial derivatives over each of the input parameters. More precisely given $f: \mathbb{R}^n \to \mathbb{R}$ and a vector (point) $x^T = (x_1, x_2, \dots, x_n)$

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \dots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix}$$

Alternately, we can express $\nabla f(x)$ as a linear combination of the standard basis for R^n, i.e.,

$$\nabla f(x) = \frac{\partial f}{\partial x_1}(x)\, e_1 + \frac{\partial f}{\partial x_2}(x)\, e_2 + \cdots + \frac{\partial f}{\partial x_n}(x)\, e_n$$

For example, if $f(x_1, x_2, x_3) = x_1^2 + x_1 x_2 + x_2^2 + x_2 x_3 + x_3^2$, then

$$\nabla f(x) = \begin{bmatrix} 2x_1 + x_2 \\ x_1 + 2x_2 + x_3 \\ x_2 + 2x_3 \end{bmatrix}$$

Using the standard basis for $\mathbb{R}^n$, we can write the above as

$$\nabla f(x) = (2x_1 + x_2)e_1 + (x_1 + 2x_2 + x_3)e_2 + (x_2 + 2x_3)e_3$$

The input point $(0,0,0)$ is the only stationary point.

4.3 Derivatives of Vector Function

A vector function $f: \mathbb{R}^m \to \mathbb{R}^n$ has a vector input and a vector output, i.e.,

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \dots \\ f_n(x) \end{bmatrix}, x \in \mathbb{R}^m$$

The derivative of f at a point x is represented by something known as the Jacobian matrix (or just "Jacobian"), denoted $Df(x)$ or $J_f(x)$ and expressed as follows:

$$J_f(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_m} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_m} \end{bmatrix}$$

For example, consider the vector function $f(x_1, x_2, x_3) = (5x_1x_2x_3, x_1^2 + x_3^2, x_2^3)$ at the point $x = (x_1, x_2, x_3)$. The associated Jacobian is

$$J_f(x) = \begin{bmatrix} 5x_2x_3 & 5x_1x_3 & 5x_1x_2 \\ 2x_1 & 0 & 2x_3 \\ 0 & 3x_2^2 & 0 \end{bmatrix}$$

Theorem 77. (Inverse Function Theorem for Vector Functions) Let $f: \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable function on an open set $U \subseteq \mathbb{R}^n$, and $x \in U$ be such that $det\ (J_f(x)) \neq 0$. Then there exists an open neighborhood $V \subseteq U$ of x and an open neighborhood $W \subseteq \mathbb{R}^n$ of $f(x)$ such that

- ***$f: V \to W$ is a bijection***
- ***The inverse function f^{-1} at $f(x)$ exists and is continuously differentiable.***
- ***The Jacobian matrix of f^{-1} at $f(x)$ is the matrix inverse of the Jacobian matrix of f at x, i.e., $J_{f^{-1}}(f(x)) = \left(J_f(x)\right)^{-1}$.***

Proof: See the article "The Implicit and the Inverse Function theorems: easy proofs" [37]. ■

4.4 Dual Numbers

In what follows, we will use the concept of **dual numbers** and vectors in our definition of derivatives as linear operators. Dual numbers are also used in a type of computer-based differentiation known as automatic differentiation.

Dual numbers are formed as pairs of real numbers with addition being component-wise and multiplication defined by the formula given below. More precisely, given dual numbers (a, b) and (c, d), addition/subtraction and multiplication are defined as follows

$$(a, b) \pm (c, d) = (a \pm c, b \pm d)$$

$$(a, b)(c, d) = (ac, ad + bc)$$

The dual number $\varepsilon = (0,1)$ has the property $\varepsilon^2 = 0$.

Alternately, we can express dual numbers in the form

$$a + b\varepsilon, \qquad a, b \in \mathbb{R}, \qquad \varepsilon^2 = 0$$

If we multiply $(a + b\varepsilon)(c + \varepsilon d)$ we get $ac + (ad + bc)\varepsilon + bd\varepsilon^2 = ac + (ad + bc)\varepsilon$ which better motivates the definition of multiplication provided above.

Division is also possible, i.e.,

$$\frac{a + b\varepsilon}{c + d\varepsilon} = \frac{a + b\varepsilon}{c + d\varepsilon} \cdot \frac{c - d\varepsilon}{c - d\varepsilon} = \frac{acz + (bc - ad)\varepsilon - bd\varepsilon^2}{c^2 - d^2\varepsilon^2} = \frac{ac + (bc - ad)\varepsilon}{c^2}, \qquad c \neq 0$$

. . .

Consider a differentiable function $f: \mathbb{R} \to \mathbb{R}$. We want to evaluate f at the dual number $x + b\epsilon$. The Taylor series of f around a point x for a perturbation h is

$$f(x + h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \frac{f'''(x)}{3!}h^3 + \cdots$$

Next, substitute $h = b\varepsilon$ and note that since $\varepsilon^2 = 0$, it follows the $\varepsilon^k = 0, k \geq 2$.

$$f(x + b\varepsilon) = f(x) + f'(x)b\varepsilon + \frac{f''(x)}{2!}(b\varepsilon)^2 + \frac{f'''(x)}{3!}(b\varepsilon)^3 + \cdots = f(x) + f'(x)b\varepsilon$$

If we let $b = 1$, we obtain

$$f(x + \varepsilon) = f(x) + f'(x)\varepsilon$$

$$f(x + \varepsilon) - f(x) = f'(x)\varepsilon$$

This is analogous to the differential $df = f'(x)dx \approx f(x + dx) - f(x)$ from calculus. Here, ε serves as an algebraic representation of an infinitesimal perturbation, similar to dx, but within the rigorous framework of dual numbers. The nilpotency of ε ensures that the Taylor series truncates exactly after the linear term, providing an exact expression for the derivative without higher-order terms.

The dual number concept can be extended to vectors. Consider

$$v = a + b\varepsilon,\ \ u = c + d\varepsilon,\ \ a, b, c, d, u, v \in \mathbb{R}^n,\ \ \varepsilon^2 = 0$$

Addition/subtraction and multiplication (dot product) are defined as follows:

$$v \pm u = (a \pm c) + (b \pm d)\varepsilon$$
$$v \cdot u = (a \cdot c) + (a \cdot d + b \cdot c)\varepsilon$$

A similar extension can be made for matrices, i.e.,

$A = A_0 + A_1\varepsilon$ where A_0 and A_1 are matrices of the same size andε is a formal symbol satisfying $\varepsilon^2 = 0$. A_0 is sometimes called the "real part" (or standard part) and A_1 is known as the "dual part" (or infinitesimal perturbation).

For dual matrices $A = A_0 + A_1\varepsilon$ and $B = B_0 + B_1\varepsilon$:

Addition: $A + B = (A_0 + B_0) + (A_1 + B_1)\varepsilon$

Scalar multiplication : $cA = cA_0 + cA_1\varepsilon$

Matrix multiplication (if dimensions allow): $AB = A_0B_0 + (A_0B_1 + A_1B_0)\varepsilon$

4.5 Derivatives as Linear Operators

[**Author's Remark**: The gradient and Jacobian are what I think of as interior techniques. These techniques are more appropriate when we are given the internal details of the scalars, vectors and matrices that define the input and output of a particular function. The examples in the previous two subsections fit that category. In other cases, e.g., $f(x) = x^TAx$, we could grind through an analysis to determine the gradient. This would involve expanding x^TAx, taking the partial derivatives to form the gradient, and then looking for a possible pattern at the level of vectors and matrices. In this section, we will discuss an approach that allows one to determine derivatives by manipulating vectors and matrices (rather than their internals). I think of this as a set of external techniques.]

4.5.1 Single-variable Calculus

In single-variable calculus, the derivative of a function $f(x)$ is defined as the slope of the tangent line at the point $(x, f(x))$. Further, $f'(x)$ can be used to approximate a change in $f(x)$ (denoted by Δf) relative to a small change in x (denoted by Δx), i.e.,

$$\Delta f = f(x + \Delta x) - f(x) = f'(x)\Delta x + o(\Delta x)$$

The notation $o(\Delta x)$ denotes any function whose magnitude shrinks much faster than $|\Delta x|$ as $\Delta x \to 0$. For example, $|\Delta x|^\alpha$ is $o(\Delta x)$ for $\alpha > 1$. More precisely,

$$\lim_{\Delta x \to 0} \frac{|\Delta x|^\alpha}{|\Delta x|} = \lim_{\Delta x \to 0} |\Delta x|^{\alpha - 1} = 0$$

As $\Delta x \to 0, o(\Delta x)$ is negligible compared to the linear term $f'(x)\Delta x$.

In differential notation, the change in the function is approximated by the differential:

$$df \equiv f'(x)dx$$

which represents the linear part of the change $f(x+dx)-f(x)$. The full change is

$$f(x+dx)-f(x) = df + o(dx)$$
$$= f'(x)dx + o(dx)$$

However, in the differential notation, we neglect the $o(dx)$ term as it vanishes for infinitesimally small dx. With this caveat mind, we take $f'(x)$ in the context of the following equation as our definition of the derivative of f at x:

$$Df \equiv f(x+dx)-f(x) \approx df = f'(x)dx \quad (1)$$

In this formulation, we avoid dividing by dx, as in done in the typical definition of derivative, i.e., $\frac{dy}{dx} = f'(x)$. This approach allows for an easier extension to differential calculus for vectors and matrices where dx is a vector (by which we cannot divide).

In terms of dual numbers, we can think of $\varepsilon = dx$, and $x+dx$ as the dual number $x+\varepsilon$. With these identifications in mind, we can restate equation (1) as $f(x+\varepsilon)-f(x) = f'(x)\varepsilon$.

From the perspective of linear algebra, the derivative of $f(x)$ can be viewed as linear transformation applied to dx, i.e.,

$$Df = f(x+dx)-f(x) = f'(x)[dx] \quad (2)$$

We have that

$$f'(x)[adx_1 + bdx_2] = af'(x)[dx_1] + bf'(x)[dx_2]$$

More generally, we can view $f'(x)$ as a linear transformation that takes an input direction v and gives an output vector $f'(x)[v]$, which can be interpreted as a directional derivative (as discussed below). When the direction v is an infinitesimal dx, the output $f'(x)[dx] = Df$ is the differential of f, i.e., the infinitesimal change in f corresponding to dx.

(In what follows, we will usually write $f'(x)dx$ and drop the brackets.)

. . .

Example 1. Find the derivative of $f(x) = x^3$ using the dual number formulation of equation (1). We have that

$$Df = f(x+\varepsilon)-f(x) = (x+\varepsilon)^3 - x^3 = x^3 + 3x^2\varepsilon + 3x\varepsilon^2 + \varepsilon^3 - x^3 = 3x^2\varepsilon$$

In the above, note that all powers 2 and above of ϵ are 0. The term multiplied times ε on the right of the equation is, by our definition, $f'(x) = 3x^2$ which we already know is the correct answer from basic calculus.

Example 2. Find the derivative of $f(x) = \sin x$ using the dual number formulation of equation (1). Using the Taylor series for $\sin x$ and noting $\varepsilon^n = 0, n \geq 2$, we have that

$$Df = \sin(x+\varepsilon) - \sin x$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n (x+\varepsilon)^{2n+1}}{(2n+1)!} - \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$$

$$= (x+\varepsilon) - \frac{(x-\varepsilon)^3}{3!} + \frac{(x-\varepsilon)^5}{5!} \mp \cdots - \left(x + \frac{x^3}{3!} - \frac{x^5}{5!} \pm \cdots \right)$$

$$= \varepsilon - \frac{x^2}{2!}\varepsilon + \frac{x^4}{4!}\varepsilon - \frac{x^6}{6!}\varepsilon \pm \cdots = (\cos x)\varepsilon$$

So, $f'(x) = \cos x$.

Another approach is to use the double angle formula for the sine function, i.e.,

$$Df = \sin(x+\varepsilon) - \sin x = \sin x \cos \varepsilon + \cos x \sin \varepsilon - \sin x$$

$$= \sin x + (\cos x)\varepsilon - \sin x = (\cos x)\varepsilon$$

In the above, we used the Taylor series for $\sin x$ and $\cos x$, and the fact that $\varepsilon^n = 0, n \geq 2$ to determine that $\cos \varepsilon = 1$ and $\sin \varepsilon = \varepsilon$.

4.5.2 Directional Derivatives

It is possible to interpret the linear operator view of a derivative as a directional derivative.

Consider a function $f(x)$ where x is a vector. The directional derivative at x in the direction of vector v is defined as:

$$\nabla_v f(x) = \lim_{\Delta a \to 0} \frac{f(x + (\Delta a)v) - f(x)}{(\Delta a)\|v\|} = \frac{1}{\|v\|} \frac{\partial}{\partial a} f(x + av)|_{a=0}, \qquad a \in \mathbb{R}$$

The above formula further simplifies if we take v to be a unit vector, i.e., $\|v\| = 1$.

On the other hand, from the linear operator definition, we have that

$$f(x + (da)v) - f(x) = f'(x)[(da)v] = (da)f'(x)[v]$$

$$\Rightarrow \frac{f(x + (da)v) - f(x)}{da} = f'(x)[v]$$

In the above, we were able to factor out da since $f'(x)$ is a linear operator (transformation) and da is a scalar. Comparing the two expressions above, we can conclude that

$$\frac{\partial}{\partial a} f(x + av)|_{a=0} = f'(x)[v]$$

Further, we have learned that it is possible (and valid) to input an arbitrary non-infinitesimal vector v into $f'(x)[v]$. While the result is not df, it is a directional derivative.

4.6 Fréchet Derivative

In the following subsection, we will extend the definition of the differential in Equation (1) to functions that take a vector input. In support of this goal, we need to use something known as the Fréchet derivative [38].

Let V and W be normed vector space, and $U \subseteq V$ be an open subset of V. A function $f: U \to W$ is said to be **Fréchet differentiable** at $x \in U$ if there exists a bounded linear operator $A: V \to W$ such that

$$\lim_{\|dx\|_V \to 0} \frac{\|f(x+dx) - f(x) - Adx\|_W}{\|dx\|_V} = 0$$

The above limit is meant in the sense of a limit of a function defined on a metric space, with V and W as the two metric spaces. Equivalently (and the form that we will use), the following first-order expansion holds:

$$df = f(x+dx) - f(x) = Adx + o(dx) \quad (3)$$

Most importantly, if such an linear operator A exists, it is unique, and we write $f'(x) = A$ and refer to it as the Fréchet derivative of f at x.

The above is more general than what we will need. For our purposes, the vector spaces V and W will be $\mathbb{R}^n$ (but not necessarily the same value of n for both V and W).

The Fréchet derivative obeys the following properties:

- $(cf)'(x) = cf'(x)$, c is a scalar
- $(f+g)'(x) = f'(x) + g'(x)$
- $(fg)'(x) = f(x)\, g'(x) + f'(x)\, g(x)$

The chain rule holds true in this context. If $f: V \to Y$ is differentiable at $x \in U$ and $g: Y \to W$ is differentiable at $y = f(x)$, then the composition $g \circ f$ is differentiable at x and the derivative is the composition of the derivatives, i.e.,

$$(g \circ f)'(x) = g'\big(f(x)\big) \circ f'(x)$$

The product rule (list above) if valid if f is a scalar-valued or matrix-valued function, i.e., its output is either a scalar or a matrix. The product rule does not hold in general for vector-valued functions. In fact, the product of two vectors is not necessarily defined in many vector spaces.

The same properties hold for the differential df.

4.7 Scalar-valued Functions

If we take U to be $\mathbb{R}^n$ and V to be $\mathbb{R}$ in the statement of Fréchet differentiable, and replace dx with the dual number εdx in equation (3), resulting in $o(\varepsilon dx)$ going to zero, then we have

$$df = f(x + \varepsilon\, dx) - f(x) = f'(x)(\varepsilon\, dx) \quad (4)$$

In the above, df is a scalar, x and dx are $n \times 1$ column vectors, and $f'(x)$ is a $1 \times n$ row vector. We already know from the previous section on the gradient that $\nabla f(x)^T$ is a solution and we know that the Fréchet derivative is unique. In particular, we have

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \qquad dx = \begin{bmatrix} dx_1 \\ dx_2 \\ \dots \\ dx_n \end{bmatrix}, \qquad f'(x) = \nabla f(x)^T = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) & \frac{\partial f}{\partial x_2}(x) & \dots & \frac{\partial f}{\partial x_n}(x) \end{bmatrix}$$

The gradient is one approach for determining the derivative of a scalar-valued function that takes a vector as an input and outputs a scalar. Another approach is to compute $df = f(x + \varepsilon\, dx) - f(x)$.

Example 3. If $f(x) = a^T x, a, x \in \mathbb{R}^n$, a is a constant vector and x is a vector of variables, determine $df, f'(x)$ and ∇f.

Note that $f(x)$ takes a vector as input and outputs a scalar.

Using equation (4), we have

$$df = f(x + \varepsilon dx) - f(x) = a^T(x + \varepsilon dx) - a^T x = a^T(\varepsilon\, dx)$$

So, $f'(x) = a^T$ and $\nabla f = f'(x)^T = a$.

Example 4. If $f(x) = x^T Ax, x \in \mathbb{R}^n$ and A is a constant $n \times n$ matrix, determine $df, f'(x)$ and ∇f.

As with the previous example, f takes a vector as input and outputs a scalar.

Using equation (4), we have

$$df = f(x + \varepsilon\, dx) - f(x) = (x + \varepsilon\, dx)^T A(x + \varepsilon\, dx) - x^T Ax$$

$$= (x^T + \varepsilon dx^T)A(x + \varepsilon dx) - x^T Ax = x^T Ax + \boldsymbol{\varepsilon x^T A dx} + \varepsilon \boldsymbol{dx^T Ax} + \varepsilon^2 dx^T Adx - x^T Ax$$

$$= \varepsilon x^T Adx + \varepsilon dx^T Ax = \varepsilon x^T Adx + \varepsilon(dx^T Ax)^T = \varepsilon x^T Adx + \varepsilon x^T A^T dx$$

$$= x^T(A + A^T)(\varepsilon\, dx)$$

In the above, note that $\varepsilon^2 = 0$, and that $dx^T Ax$ is a scalar and thus, equals its transpose. Further, for a square matrix, it is always true that $A + A^T = A^T + A$.

So, $f'(x) = x^T(A + A^T)$ and $\nabla f = f'(x)^T = (A + A^T)x$.

If we used the definition of the gradient, we would have needed to expand all the vectors, take partial derivatives and then reformulate in terms of vectors.

Example 5. If $f(X) = tr(X)$ where X is an $n \times n$ matrix of real variables, determine df and $f'(X)$.

Using Equation (3), we have

$$df = tr(X + dX) - tr(X) = tr(dX)$$

To determine $f'(X)$, we will compute each partial derivative directly. Since $f(X) = \sum_{i=1}^{n} x_{ii}$, the partial derivative $\frac{\partial f}{\partial x_{ij}}$ is equal to 1 if $i = j$ (the diagonal elements) and 0 if $i \neq j$ (the off-diagonal elements). Thus, the full set of partial derivatives of f can be represented by the identity matrix I. This is not exactly the gradient since the gradient is for functions that map from $\mathbb{R}^n$ to $\mathbb{R}$ and $tr(X)$ maps from the vector space of $n \times n$ matrices to $\mathbb{R}$. So, what we have is sort of a 2-dimensional gradient. It is also not the Jacobian, since the Jacobian is defined for functions that map from $\mathbb{R}^m$ to $\mathbb{R}^n$. This type of 2-dimensional gradient is needed in general when we seek to find the derivative of a function that takes a matrix as input and outputs a scalar.

Example 6. If $f(X) = \det X$ where X is an $n \times n$ matrix of real variables, determine df and $f'(X)$.

The differential of f can be expressed as the sum of the differentials $(df)_{ij}$, i.e.,

$$df = \sum_{i=1}^{n}\sum_{j=1}^{n}(df)_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial f}{\partial x_{ij}}dx_{ij}$$

where $\frac{\partial f}{\partial x_{ij}}$ is the partial derivative of f with respect to element (i, j) of X, and dx_{ij} is the differential of x_{ij}.

As was the case with the previous example, we need a 2-dimensional gradient to represent $f'(X)$ since f maps from an $n \times n$ matrix to a scalar.

For the determinant function, the partial derivative $\frac{\partial \det(X)}{\partial x_{ij}}$ is equal to cofactor C_{ij} of matrix X (this follows from the expansion of the $\det(X)$ in terms of cofactors, see Section 3.9). Since the adjugate matrix $adj(X)$ is the transpose of the cofactor matric C, we have

$$\frac{\partial \det(X)}{\partial x_{ij}} = C_{ij} = \big(adj(X)\big)_{ji}$$

Thus, we can write the differential df as

$$df = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial \det(X)}{\partial x_{ij}}\,dx_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n}C_{ij}\,dx_{ij} = tr(adj(X)\,dX)$$

(The above follows from the formula $tr(A^T B) = \sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}b_{ij}$ for $m \times n$ matrices A and B. This is known as the **Frobenius inner product** of A and B.)

Further, since $adj(X) = \det(X)\,X^{-1}$ when X is invertible, we can also write

$$df = tr(\det(X)\,X^{-1}dX) = \det(X)\,tr(X^{-1}dX)$$

The derivative $f'(X)$ is the matrix that represents the linear transformation df with respect to dX. In matrix calculus, when f is a scalar function of a matrix X, the differential is often expressed as $df = tr(\nabla f^T\, dX)$, where $\nabla f = f'(X)$ is the 2-dimension gradient of f with respect to X. So, from the first expression for df above, we have that $f'(X) = adj(X)^T = C(X)$, i.e., the cofactor matrix for X. When X is invertible, we also have $f'(X) = \det(X)\,(X^{-1})^T$.

4.8 Vector-valued Functions

If, in Equation (3), we let U be $\mathbb{R}^m$ and V be $\mathbb{R}^n$, and replace dx with the dual number εdx, the $o(\varepsilon dx)$ goes to zero, and we have

$$df = f(x + \varepsilon\, dx) - f(x) = f'(x)(\varepsilon\, dx) \quad (5)$$

In the above, df is an $n \times 1$ column vector, x and dx are $m \times 1$ column vectors, and $f'(x)$ is an $n \times m$ matrix. We already know from the previous section on the Jacobian that $J_f(x)$ is a solution and we know that the Fréchet derivative is unique. In particular, we have

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_m \end{bmatrix}, \qquad dx = \begin{bmatrix} dx_1 \\ dx_2 \\ \cdots \\ dx_m \end{bmatrix}, \qquad f'(x) = J_f(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_m} \end{bmatrix}$$

As we shall see in several examples, the approach based on the evaluation of df can be much easier than the Jacobian approach.

Example 7. If $f(x) = Ax$ where $x \in \mathbb{R}^n$ is a vector of variables, and A is an $m \times n$ constant matrix, then determine df and $f'(x)$.

Using Equation (5), we have

$$df = f(x + \varepsilon\, dx) - f(x) = A(x + \varepsilon\, dx) - Ax = A(\varepsilon\, dx)$$

So, $f'(x) = A$.

Example 8. What if we change the previous example slightly and take $f(X) = Xa$ where X is an $m \times n$ matrix of variables, and $a \in \mathbb{R}^n$ is a constant vector? In this case, we have

$$df = f(X + \varepsilon\, dX) - f(X) = (X + \varepsilon\, dX)a - Xa = \varepsilon\, (dX)a$$

Here, dX is an $m \times n$ matrix of differentials. Note that dX and a do not commute. So, it appears we are a bit stuck in terms of extracting $f'(X)$ from our formula for df. However, there is a solution. From Theorem 72, we have that $vec(dXa) = vec(I(dX)a) = (a^T \otimes I)vec(\varepsilon\, dX)$. Thus,

$$vec(df) = (a^T \otimes I)vec(dX)$$

It then follows that $f'(X) = a^T \otimes I$ when expressed as a matrix acting on $vec(dX)$.

4.9 Matrix-valued Functions

We can again extend our formula for df in the case of functions that output a matrix. Depending on the input of $f(x)$, the dx term in the formula $df = f(x + \varepsilon\, dx) - f(x) = f'(x)(\varepsilon\, dx)$ can be a scalar, vector or matrix. Further, we can use the dual number approach, or we can just go with the form $df = f(x + dx) - f(x) = f'(x)(dx)$. In the latter approach, we need to delete $o(dx)$ terms such as $(dx)^2$.

Example 9. Determine df and $f'(X)$ if $f(X) = X^3$ with X being an $n \times n$ matrix of variables.

Let's first determine dg for $g(X) = X^2$. Using the product rule, we have

$$dg = X\, dX + dX\, X$$

Next apply the product rule to $X^3 = X^2 X$.

$$df = X^2\, dX + d(X^2)\, X = X^2\, dX + (X\, dX + dX\, X)X$$

$$= X^2\, dX + X\, dX\, X + dX\, X^2$$

We cannot write the above as $3X^2\, dX$ since we are dealing with matrices and the commutativity law does not hold for matrices. However, we can write df as $X^2\, dX\, I + X\, dX\, X + I\, dX\, X^2$ and then apply Theorem 72 to each of the terms to get

$$vec(df) = (I \otimes X^2 + X^T \otimes X + (X^2)^T \otimes I)\, vec(dX)$$

Thus, $f'(X) = (I \otimes X^2 + X^T \otimes X + (X^2)^T \otimes I)$ when expressed as a matrix acting on $vec(dX)$.

More generally, if $f(X) = X^k$ for an $n \times n$ matrix of variable and $k \in \{1,2,3, \dots\}$, then we have

$$df = \sum_{i=0}^{k-1} X^{k-1-i}(dX)X^i$$

$$vec(df) = \left(\sum_{i=0}^{k-1} \left(X^i\right)^T \otimes X^{k-1-i}\right) vec(dX)$$

So, as a matrix acting on $vec(dX)$,

$$f'(X) = \sum_{i=0}^{k-1} \left(X^i\right)^T \otimes X^{k-1-i}$$

If X and dX commute, the differential simplifies to $df = kX^{k-1}dX$.

Example 10. Determine df and $f'(X)$ if $f(X) = X^{-1}$ with X being an invertible $n \times n$ matrix of variables.

Since $XX^{-1} = I$, we have (using the product rule)

$$0 = d(I) = d(XX^{-1}) = dX\, X^{-1} + X\, d(X^{-1})$$

$$d(X^{-1}) = -X^{-1}\, dX\, X^{-1}$$

Using Theorem 72, we can derive a vectorized version of the above formula and then extract $f'(X)$.

$$vec\big(d(X^{-1})\big) = -((X^{-1})^T \otimes X^{-1})vec(dX)$$

$$\Rightarrow f'(X) = -((X^{-1})^T \otimes X^{-1})$$

when expressed as a matrix acting on $vec(dX)$.

4.10 Hessian Matrices and Second Derivatives

4.10.1 Second Derivatives of Scalar-valued Functions

The **Hessian matrix** is a square matrix of second-order partial derivatives of a scalar-valued function. It describes the local curvature of a function of many variables. The Hessian matrix was developed in the 19th century by the German mathematician Ludwig Otto Hesse and later named after him. There are various notations for the Hessian matrix. In this book, we will use the notations $\nabla^2 f$ or H_f. [Many sources will just write H for the Hessian but this allows for ambiguity since this notation does not reference the function for which the Hessian is applied.]

Let $f: \mathbb{R}^n \to \mathbb{R}$ be a function taking a vector $x \in \mathbb{R}^n$ as input and giving output $f(x) \in \mathbb{R}$. If all the second-order partial derivatives of f exist, then the Hessian matrix $\nabla^2 f$ is an $n \times n$ matrix, defined and arranged as follows:

$$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

If the second partial derivatives are all continuous, the Hessian matrix is a symmetric matrix by the symmetry of second derivatives [39], i.e., $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ and thus, $H_f = H_f^T$.

The Hessian matrix of a function f is the Jacobian matrix of the gradient of the function f, i.e., J_f.

For example, , if $f(x_1, x_2, x_3) = x_1^3 + x_1 x_2 + x_2^3 + x_2 x_3 + x_3^3$, then

$$\nabla f(x) = \begin{bmatrix} 3x_1^2 + x_2 \\ x_1 + 3x_2^2 + x_3 \\ x_2 + 3x_3^2 \end{bmatrix}, \qquad \nabla^2 f(x) = \begin{bmatrix} 6x_1 & 1 & 0 \\ 1 & 6x_2 & 1 \\ 0 & 1 & 6x_3 \end{bmatrix}$$

4.10.2 Second Derivatives of Vector-valued Functions

Consider functions that map from $\mathbb{R}^m$ to $\mathbb{R}^n$ with both m and n greater than 1. Various approaches are possible. We will use a specific example to illustrate each approach, i.e.,

$$f(x,y) = \begin{bmatrix} f_1(x,y) \\ f_2(x,y) \end{bmatrix} = \begin{bmatrix} x^2 y \\ y\sin(x) \end{bmatrix}$$

The first approach is something known as a **third-order tensor** (tensors are the topic of Section 6 where we will cover the topic in more detail). For $f\colon \mathbb{R}^m \to \mathbb{R}^n$, the second derivative can be expressed with the following third-order tensor:

$$T_{ijk} = \frac{\partial^2 f_i}{\partial x_j \partial x_k}, \qquad i = 1,2,\dots,n, \qquad j,k = 1,2,\dots,m$$

For our example, the second derivative is a $2 \times 2 \times 2$ tensor.

For f_1, we have

$$\frac{\partial^2 f_1}{\partial x^2} = 2y, \qquad \frac{\partial^2 f_1}{\partial x \partial y} = 2x, \qquad \frac{\partial^2 f_1}{\partial y \partial x} = 2x, \qquad \frac{\partial^2 f_1}{\partial y^2} = 0$$

So, the slice $T_{1,:,:}$ is

$$\begin{bmatrix} 2y & 2x \\ 2x & 0 \end{bmatrix}$$

For f_2, we have

$$\frac{\partial^2 f_2}{\partial x^2} = -y\sin(x), \qquad \frac{\partial^2 f_2}{\partial x \partial y} = \cos(x), \qquad \frac{\partial^2 f_2}{\partial y \partial x} = \cos(x), \qquad \frac{\partial^2 f_2}{\partial y^2} = 0$$

So, the slice $T_{2,:,:}$ is

$$\begin{bmatrix} -y\sin(x) & \cos(x) \\ \cos(x) & 0 \end{bmatrix}$$

The final tensor representation is

$$T = \left[\begin{bmatrix} 2y & 2x \\ 2x & 0 \end{bmatrix}, \begin{bmatrix} -y\sin(x) & \cos(x) \\ \cos(x) & 0 \end{bmatrix} \right]$$

. . .

The next approach is to compute an **array of Hessians**. In this approach we determine the Hessian for each component f_i of f. We then have $f'' = (H_{f_1}, H_{f_2}, \dots, H_{f_n})$.

For our example, we have

$$H_{f_1} = \begin{bmatrix} 2y & 2x \\ 2x & 0 \end{bmatrix}, \qquad H_{f_2} = \begin{bmatrix} -y\sin(x) & \cos(x) \\ \cos(x) & 0 \end{bmatrix}, \qquad f'' = (H_{f_1}, H_{f_2})$$

. . .

Another approach is to compute the first derivative (i.e., Jacobian), **flatten the Jacobian** into a vector and then compute the Jacobian again to get the second derivative of the original function.

For example, the first derivative is

$$J_f = \begin{bmatrix} 2xy & x^2 \\ y\cos(x) & \sin(x) \end{bmatrix}$$

We then flatten (i.e., vectorize) J_f, i.e.,

$$vec(J_f) = \begin{bmatrix} 2xy \\ y\cos(x) \\ x^2 \\ \sin(x) \end{bmatrix}$$

To get f'', we compute the Jacobian of $vec(J_f)$ with respect to x and y, i.e.,

$$f'' = J_{vec(J_f)} = \begin{bmatrix} 2y & 2x \\ -y\sin(x) & \cos(x) \\ 2x & 0 \\ \cos(x) & 0 \end{bmatrix}$$

So, we get the same information as in the previous two approaches but in a different arrangement.

. . .

[**Author's Remark**: Recall my remarks at the beginning of Section 4.5. The above approaches for the second derivative are interior techniques. As such, these techniques are more appropriate when we are given the internal details of the scalars, vectors and matrices that define the input and output of a particular function. In the following subsection, we will discuss an approach involving bilinear forms that is more appropriate when working at the level of vectors and matrices.]

4.10.3 Second Derivatives via Differentials

If presented with a function in terms of vectors and matrices (as opposed to the internal detail as in the examples from the previous subsection), it is better to use the differential approach. The idea is to apply the differential approach to determine the first derivative, and then a second time to determine the second derivative.

Example 11. If $f(x) = x^T Ax, x \in \mathbb{R}^n$ and A is a constant $n \times n$ matrix, determine $f''(x)$.

We already determined $f'(x)$ in Example 4, i.e., $f'(x) = x^T(A + A^T)$. Let $g(x) = f'(x)$ and apply our definition of differential to g.

$$dg = g(x + dx) - g(x) = (x + dx)^T(A + A^T) - x^T(A + A^T)$$

$$= x^T(A + A^T) + dx^T(A + A^T) - x^T(A + A^T)$$

$$= dx^T(A + A^T)$$

There is an issue here since, by definition, $dg = g'(x)dx$ and the above has the dx term on the left. The solution is the $dx^T(A + A^T)$ is a row vector and $(A + A^T)dx$ is a column vector, but they have the same components. So, we conclude that $g'(x) = A + A^T \Rightarrow f''(x) = A + A^T$.

Another approach is to note that $d^2f = d(df)$. So, we can find df and then take the differential a second time. From Example 4, we already have that $df = x^T(A + A^T)dx$. Using the product rule, we have

$$d^2f(x) = d\big(df(x)\big) = d(x^T(A + A^T)dx) = dx^T(A + A^T)dx + x^T(A + A^T)d^2x$$

$$= dx^T(A + A^T)dx$$

Assuming x is not depending on another variable, $d^2x = 0$.

We still have the task of extracting $f''(x)$ from the above expression. To that end, let $f(x)$ be any function that has a vector input and outputs a scalar (the problem at hand is one such example). By definition, we have that $d^2f = d(df) = d(\nabla f(x)^T dx)$. Here, dx is treated as a constant (since it is independent of x), and so, we can move dx out of the differential, i.e., $d\big(\nabla f(x)\big)^T dx$.

The differential of the gradient $\nabla f(x)$ is $d(\nabla f(x)) = \nabla^2 f(x)dx$ where $\nabla^2 f(x) = H$ is the Hessian matrix. Substituting back:

$$d^2f = (H\ dx)^T dx = dx^T H dx$$

Applying this to our original problem, it must be the $H = A + A^T$ is $f''(x)$.

Example 12. Given $f(X) = X^3$ where X is an $n \times n$ matrix of variables, find d^2f and $f''(X)$.

In a previous exercise, we found that

$$df = X^2\ dX + X\ dX\ X + dX\ X^2$$

To compute d^2f, we will also need the differential of X^2 which is $X(dX) + (dX)X$.

Using the above result for df, we have

$$d^2f = d(df) = (\ X(dX) + (dX)X)dX + X^2(d^2X)$$

$$+dX(dX\ X) + X((d^2X)X + (dX)dX))$$

$$+(d^2X)X^2 + dX(X(dX) + (dX)X$$

In the above, the d^2X terms are zero, giving us

$$d^2f = X(dX)^2 + (dX)X(dX) + (dX)^2X + X(dX)^2 + (dX)X(dX) + (dX)^2X$$

$$= 2X(dX)^2 + 2(dX)X(dX) + 2(dX)^2X$$

$$= 2(X(dX)^2 + (dX)X(dX) + (dX)^2X)$$

If we eliminate the $o(dX)$ terms, then we have

$$d^2f = 2(X(dX)^2 + (dX)X(dX) + (dX)^2X)$$

So, viewing $f''(X)$ as a bilinear operator, we have

$$f''(X)[dX, dX] = 2(X(dX)^2 + (dX)X(dX) + (dX)^2X)$$

As a further exercise, we can vectorize d^2f. In what follows, we vectorize each of the terms in $2(X(dX)^2 + (dX)X(dX) + (dX)^2X)$ (applying Theorem 72 multiple times) and then collect terms to get the final solution.

For the first term, we have

$$vec(X(dX)^2) = vec(X(dX)^2 I) = (I \otimes X)vec(dX\ dX)$$

$$vec(dX\ dX) = vec(dX\ dX\ I) = (I \otimes dX)vec(dX)$$

So, $vec(X(dX)^2) = (I \otimes X)(I \otimes dX)vec(dX)$.

For the second term, we have

$$vec(dX\ X\ dX) = vec\big((dX\ X)\ dX\ I\big) = \big(I \otimes (dX\ X)\big)vec(dX)$$

For the third term, we have

$$vec((dX)^2X) = vec(I\ (dX)^2\ X) = (X^T \otimes I)vec(dX\ dX)$$

$$= (X^T \otimes I)(I \otimes dX)vec(dX)$$

Putting the above result together yields

$$vec(d^2f) = 2\Big((I \otimes X)(I \otimes dX) + \big(I \otimes (dX\ X)\big) + (X^T \otimes I)(I \otimes dX)\Big)\, vec(dX)$$

Using the identity $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$, the above simplifies to

$$vec(d^2f) = 2\big[(I \otimes XdX) + \big(I \otimes (dX\ X)\big) + (X^T \otimes dX)\big]vec(dX)$$

4.11 Further Study and References

There are several books and Internet sites that provide lists of known derivatives of functions involving vectors or matrices in their inputs or outputs.

- *The Matrix Reference Manual*, Mike Brookes, http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/intro.html, see the section entitled "Calculus"
- *Matrix calculus*, Wikipedia, https://en.wikipedia.org/wiki/Matrix_calculus
- *The Matrix Cookbook*, K. B. Petersen and M. S. Peders, https://archive.org/details/K_B_Petersen_and_M_S_Peders__The_Matrix_Cookbook

MIT offers material (videos and lecture notes) from their course "Matrix Calculus for Machine Learning and Beyond", see https://ocw.mit.edu/courses/18-s096-matrix-calculus-for-machine-learning-and-beyond-january-iap-2023/

"A gentle introduction to matrix calculus" by Jan R. Magnus provides a very technical (and not easy to read) introduction to matrix calculus, see https://www.sciencedirect.com/science/article/pii/S0304407624002070. This introduction is extracted from an extensive book on the topic, i.e.,

> Magnus, J.R., Neudecker, H., 2019. *Matrix Differential Calculus with Applications in Statistics and Econometrics*, third ed. John Wiley, Chichester/New York.

5 Indical Notations

Education consists mainly of what we have unlearned.

Mark Twain

5.1 Einstein Notation

5.1.1 Definition

Einstein notation, also known as Einstein summation convention, is a notational shorthand used in mathematics and physics (especially in linear algebra, tensor calculus, and differential geometry) to simplify expressions involving summations over indexed variables. We introduce Einstein notation here because we will need it in the following section concerning tensors.

Sometimes you will see the term “indicial notation”. This is a more general concept of which Einstein notation is an example.

Albert Einstein did invent the Einstein notation. He first introduced this notational convention in 1916. This shorthand method was developed to simplify the complex mathematical equations he was working with, particularly in the context of his theory of general relativity. The notation, also known as the Einstein summation convention, eliminates the need to write the summation symbol in certain contexts. Einstein himself is said to have humorously remarked on his innovation, stating, "I have made a great discovery in mathematics; I have suppressed the summation sign."

According to this convention, when an index variable appears twice in a single term and is not otherwise defined, it implies summation of that term over all the values of the index, i.e., the expression $a^i b_i$ is shorthand for $\sum_i a^i b_i$. The upper indices are not exponents but rather indices. In our example, in this context a^i should be understood as the i^{th} component of a rather than a to the power i. [Yes, this takes some getting used to, but the convention is widely used in various areas.]

An index that is summed over is referred to as the **summation index** (that would i in the above example). It is also called a dummy index since any symbol can replace i without changing the meaning of the expression (provided that it does not collide with other index symbols in the same term), e.g., $a^i b_i$ and $a^j b_j$ represent the same summation.

An index that is not related to summation is referred to as a **free index** and should appear only once per term. For example, in the expression $a^i b_i c_j = c_j \sum_i a^i b_i$, j is a free index and as such, we can move it out from the summation.

A few more examples:

- $a^i b_i c^j d_j = \sum_i a^i b_i \sum_j c^j d_j$
- $a^i_j b^j_k$ for $j = 1,2,3$ is equivalent to $a^i_1 b^1_k + a^i_2 b^2_k + a^i_3 b^3_k$ where i and k are free indices
- $a^i_j b^j_i$ for $i = 1,2,3, j = 1,2,3$ is equivalent to

$$\sum_{i=1}^{3} \sum_{j=1}^{3} a^i_j b^j_i$$

5.1.2 Rules

The following are the rules to be followed when using Einstein notation.

1. **Implicit Summation over Repeated Indices**: If an index appears exactly twice in a single term, once as a superscript (also known as the contravariant index) and once as a subscript (also known as the covariant index), summation over that index is implied.
2. **Each Index Can Appear at Most Twice in Any Term**: An index cannot appear more than twice within a single term. If an index appears three or more times in a term, it's considered an invalid expression in Einstein notation. For example, the expression $a^i b_i c^i$ is invalid.
3. **Free Indices Must Match on Both Sides of an Equation**: Indices that are not repeated within a term are called "free indices." If an index is a free index on one side of an equation, it must also be a free index on the other side, and its position (superscript or subscript) must match. For example, $a_i = b_i + c_{ij} d^j$ is valid but $a_i = b_j$ is not valid since the free indices don't match. The expression $a_i = b^i$ is also invalid since the free index on the left is a subscript and the free index on the right is a superscript.

5.1.3 Additional Examples

If we let the subscript (or superscript) represent a row vector or a row of a matrix, and a superscript (or subscript) represent a column vector or column of a matrix, then we can represent matrix multiplication using Einstein notation.

For example, the matrix equation $Ax = b$ can be represented by $A_i^j x^j = b^i$ where A_i^j is element (i, j) in matrix A, x^j is the j^{th} element in column vector x and b^i is the i^{th} element in column vector b. Alternately, we could write this as $A_j^i x_j = b_i$ where, in this case, A_j^i is element (i, j) in matrix A. In the first approach, the subscript represents rows and the superscript represents columns. In the second approach, we invert the meaning of the subscript and superscript.

We can represent the matrix product $C = AB$ as $C_i^j = A_i^k B_k^j$ where C_i^j is element (i, j) of matrix C. Alternately, we could write this as $C_j^i = A_k^i B_j^k$ where C_j^i is element (i, j) of matrix C, with a similar interpretation of indices for A and B.

. . .

The inner product of row vector u and column vector v (assuming both have the same number of elements) can be represented as $\langle u, v \rangle = u_i v^i = \sum_i u_i v^i$.

. . .

The trace of matrix A can be represented as $A_i^i = \sum_i A_i^i$

. . .

Recall that the Kronecker delta function δ_j^i is 1 if $i = j$ and 0 if $i \neq j$, for $i = 1,2, \ldots, n$.

In terms of Einstein notation,

- $\delta_i^i = \sum_i \delta_i^i = n$ if the index i goes from 1 to n
- For $n \times 1$ vector v, $\delta_j^i v_i = \delta_j^1 v_1 + \cdots + \delta_j^j v_j + \cdots + \delta_j^n v_n = v_j$
- For matrix A, we have that $\delta_m^i A_j^m = A_j^i$. Since we can view δ as a matrix and apply it to itself, we have $\delta_m^i \delta_j^m = \delta_j^i$.
- If we take the standard basis for $\mathbb{R}^n$, i.e., $\{e^1, e^2, \ldots, e^n\}$, then $\langle e^i, e_j \rangle = \delta_j^i$. Note that e_j is the j^{th} standard basis vector but represented as a column.

5.2 Levi-Civita Symbol

The **Levi-Civita symbol** or Levi-Civita epsilon represents a collection of numbers defined from the sign of a permutation of the natural numbers $1, 2, \ldots, n$, for some positive integer n. It is named after the Italian mathematician and physicist Tullio Levi-Civita. Other names include the permutation symbol, antisymmetric symbol, or alternating symbol, which refer to its antisymmetric property and definition in terms of permutations.

In n dimensions, the Levi-Civita symbol is defined as follows:

$$\epsilon_{a_1 a_2 \ldots a_n} = \begin{cases} +1, & \text{if } (a_1\ a_2 \ldots a_n) \text{ is an even permutation} \\ -1, & \text{if } (a_1\ a_2 \ldots a_n) \text{ is an odd permutation} \\ 0, & \text{otherwise} \end{cases}$$

Thus, it is the sign of the permutation in the case of a permutation, and zero otherwise.

. . .

In terms of a quick review of permutations, the notation $(a_1\ a_2 \ldots a_n)$ is a rearrangement of the integers $1,2, \ldots, n$ such that $1 \to a_1, 2 \to a_2, \ldots, n \to a_n$. This is sometime written in a longer form, i.e.,

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ a_1 & a_2 & \ldots & a_n \end{pmatrix}$$

A cycle within a permutation is a sequence of mappings that return to the starting point. For example, the following permutation has two cycles, i.e., $1 \to 4 \to 6 \to 1$ and $2 \to 5 \to 3 \to 2$.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 2 & 6 & 3 & 1 \end{pmatrix}$$

We can write the above in cyclic notation as the composition of two cycles, i.e., $(1,4,6)(2,5,3)$. Notice that we separated the elements with comas to indicate a cycle. On the other hand, the above permutation (not expanded as cycles) is written as $(4\ 5\ 2\ 6\ 3\ 1)$.

Every permutation can be decomposed into a composition of cycles.

An even permutation can be obtained as the composition of an even number (and only an even number) of exchanges (called transpositions) of two elements, while an odd permutation can be obtained by (and only by) an odd number of transpositions. A cycle with only two elements is called a transposition.

For example, the following permutation has two transpositions (i.e., $1 \leftrightarrow 3$ and $2 \leftrightarrow 4$) and thus, is even.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$$

The following composition rules apply for permutations:

- the composition of two even permutations is even
- the composition of two odd permutations is even
- the composition of an odd and an even permutation is odd

From the above, it follows that

- the inverse of every even permutation is even
- the inverse of every odd permutation is odd

. . .

In the case $n = 3$, we have

$$\epsilon_{ijk} = \epsilon^{ijk} = \begin{cases} +1, & \text{if } (i\ j\ k) \text{ is } (1\ 2\ 3), (2\ 3\ 1), (3\ 1\ 2) \\ -1, & \text{if } (i\ j\ k) \text{ is } (3\ 2\ 1), (2\ 1\ 3), (1\ 3\ 2) \\ 0, & \text{otherwise} \end{cases}$$

If we permute the indices, we have the following identities:

$$\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij} = -\epsilon_{kji} = -\epsilon_{jik} = -\epsilon_{ikj}$$

. . .

In three dimensions, the Levi-Civita symbol is related to the cross product. Let $\{\boldsymbol{e_1}, \boldsymbol{e_2}, \boldsymbol{e_3}\}$ be a positively oriented orthonormal basis of a vector space V. For vectors $u = (u^1, u^2, u^3), v = (v^1, v^2, v^3) \in V$, their cross product can be represented in terms of the following determinant:

$$u \times v = \det \begin{vmatrix} \boldsymbol{e_1} & \boldsymbol{e_2} & \boldsymbol{e_2} \\ u^1 & u^2 & u^3 \\ v^1 & v^2 & v^3 \end{vmatrix} = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{ijk}\, e_i\, u^j\, v^k$$

Component i of $u \times v$ is

$$(u \times v)^i = \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{ijk}\, u^j\, v^k = \epsilon_{ijk}\, u^j\, v^k$$

The expression on the right of the above is written in terms of Einstein notation.

For n dimensions, we have the following formula that relates the Kronecker delta and Levi-Civita symbols:

$$\epsilon_{a_1 a_2 \dots a_n}\epsilon_{b_1 b_2 \dots b_n} = \det \begin{vmatrix} \delta_{b1}^{a_1} & \delta_{b2}^{a_1} & \dots & \delta_{bn}^{a_1} \\ \delta_{b1}^{a_2} & \delta_{b2}^{a_2} & \dots & \delta_{bn}^{a_2} \\ \dots & \dots & \dots & \dots \\ \delta_{b1}^{a_n} & \delta_{b2}^{a_n} & \dots & \delta_{bn}^{a_n} \end{vmatrix}$$

For a proof, see the Wikipedia article "Levi-Civita symbol" [40].

The formula can be modified for the cases $n = 2,3$, i.e.,

$$\epsilon_{ij}\epsilon_{kl} = \det \begin{vmatrix} \delta_k^i & \delta_l^i \\ \delta_k^j & \delta_l^j \end{vmatrix}, \qquad \epsilon_{ijk}\epsilon_{lmn} = \det \begin{vmatrix} \delta_l^i & \delta_m^i & \delta_n^i \\ \delta_l^j & \delta_m^j & \delta_n^j \\ \delta_l^k & \delta_m^k & \delta_n^k \end{vmatrix}$$

$$\dots$$

Theorem 78. $\boldsymbol{\epsilon^{ijk}u_i v_j w_k = det \begin{vmatrix} u^1 & u^2 & u^3 \\ v^1 & v^2 & v^3 \\ w^1 & w^2 & w^3 \end{vmatrix}}$ ***where*** $\boldsymbol{u, v, w \in \mathbb{R}^3}$.

Proof: Since ϵ^{ijk} is non-zero only when i, j, k are distinct, we can write out all $3! = 6$ permutations of $(1,2,3)$.

We have that

$$\epsilon^{ijk}u_i v_j w_k = \sum_{i,j,k} \epsilon^{ijk}u_i v_j w_k$$

$$= \epsilon^{123}u_1 v_2 w_3 + \epsilon^{132}u_1 v_3 w_2 + \epsilon^{213}u_2 v_1 w_3 + \epsilon^{231}u_2 v_3 w_1 + \epsilon^{312}u_3 v_1 w_2 + \epsilon^{321}u_3 v_2 w_1$$

$$= u_1 v_2 w_3 - u_1 v_3 w_2 - u_2 v_1 w_3 + u_2 v_3 w_1 + u_3 v_1 w_2 - u_3 v_2 w_1$$

$$= det \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

6 Tensors

Share your knowledge. It is a way to achieve immortality.

Dalai Lama XIV

6.1 Tensors as Multilinear Maps

In this subsection, we define and develop the concept of a **tensor** as a multilinear map. The reader is advised to take a second look at Sections 2.17 (Bilinear Maps) and 2.18 (Multilinear Maps) as these are examples of tensors. We will also make use of the Dual Space concept (see Section 2.19). The presentation in this subsection follows along the lines of that provided in "Comprehensive Introduction to Linear Algebra: Part III – Operators and Tensors" [41].

Let V be a finite-dimensional vector space over some field F (we will confine ourselves to the field $\mathbb{R}$ which is sufficient for applications in physics and machine learning), and let V^r denote the r-fold Cartesian product $V \times V \times \ldots \times V$. Further, let $T: V^r \to \mathbb{R}$ be a multilinear mapping. Such a mapping is known as an r-tensor on V (also known as an r-linear form on V or a multilinear form of degree r on V). The set of all r-tensors is denoted by $\mathcal{T}_r(V)$.

Addition and scalar multiplication on $\mathcal{T}_r(V)$ are defined as follows:

$$(S+T)(v_1, \ldots, v_r) = S(v_1, \ldots, v_r) + T(v_1, \ldots, v_r)$$
$$(\alpha T)(v_1, \ldots, v_r) = \alpha T(v_1, \ldots, v_r)$$

where $S, T \in \mathcal{T}_r(V), \quad \alpha \in \mathbb{R}, \quad v_i \in V, i = 1,2, \ldots, r.$

With the above operations, $\mathcal{T}_r(V)$ forms a vector space over $\mathbb{R}$.

We can extend the above definition to get a more general concept of a tensor by including support for dual spaces. In particular, we define a tensor T on V to be a multilinear map

$$T: (V^* \times \ldots \times V^*) \times (V \times \ldots \times V) \to \mathbb{R}$$

where V^* is the dual space to V, V^* is repeated s times and V is repeated r times.

The number of instances r of V is known as the **covariant order**, and the number of instances s of V^* is known as the **contravariant order** of T. A tensor of covariant order r and contravariant order s is said to be of type (r, s). The set of all tensors of type (r, s) is denoted $\mathcal{T}_r^s(V)$. With the operations defined above, $\mathcal{T}_r^s(V)$ is a vector space over $\mathbb{R}$ (which we will refer to as a **tensor space**).

A tensor of type $(0,0)$ is defined as a scalar. A tensor of type $(0,1)$ is known as a contravariant vector, and a tensor of type $(1,0)$ is known as a covariant vector (or just "covector").

In what follows, basis vectors for V will be indicated with a subscript (e.g., e_i) and basis vectors for V^* will be indicated with a superscript (e.g., ω^j). In addition, $\mathcal{T}_r^s(V)$ will be written as $\mathcal{T}_r^s$ with the vector space V and field $\mathbb{R}$ being understood.

Take $T \in \mathcal{T}_r^s$. Let $\{e_1, \ldots, e_n\}$ be a basis for V, and $\{\omega^1, \ldots, \omega^n\}$ be a basis for V^*. For $i = 1,2, \ldots, r$, define the vectors $u_i = e_j a_i^j \in V$ where $a_i^j \in \mathbb{R}$ is the j^{th} component of u_i. For $k = 1,2, \ldots, s$, define a vector $v^k = \omega^l b_l^k \in V^*$ where $b_l^k \in \mathbb{R}$ is the l^{th} component of v^k. Using the multilinearity of T and Einstein notation, we have that

$$T(v^1, \dots, v^s, u_1, \dots, u_r) = T(\omega^{l_1} b^1_{l_1}, \dots, \omega^{l_s} b^s_{l_s}, e_{j_1} a^{j_1}_1, \dots, e_{j_r} a^{j_r}_r)$$

$$= b^1_{l_1} \dots b^s_{l_s} a^{j_1}_1 \dots a^{j_r}_r \, T(\omega^{l_1}, \dots, \omega^{l_s}, e_{j_1}, \dots, e_{j_r})$$

The Einstein convention continues to apply to the second line of the above, i.e., there are $r + s$ sets of indices, each summed over n values (the dimension of V), leading to n^{r+s} terms in the sum.

The n^{r+s} scalars $T(\omega^{l_1}, \dots, \omega^{l_s}, e_{j_1}, \dots, e_{j_r})$ are known as the **components** of T relative to the bases $\{e_1, \dots, e_n\}$ and $\{\omega^1, \dots, \omega^n\}$, and are denoted by

$$T^{l_1 \dots l_s}_{j_1 \dots j_r}$$

Example 1. Consider the tensor space $\mathcal{T}^3_3$ defined by the components

$$T^{ijk}_{lmn} = \delta^l_i \delta^m_j \delta^n_k + \epsilon_{ijk} \epsilon^{lmn}$$

[There are **no** Einstein summations implied in the above. Also, the vector space V is not mentioned but we are given that V is of dimension 3.]

We can compute tensors for some specific values of the indices, e.g.,

$$T^{123}_{123} = \delta^1_1 \delta^2_2 \delta^3_3 + \epsilon_{123} \epsilon^{123} = 1 \cdot 1 \cdot 1 + 1 \cdot 1 = 2$$

$$T^{112}_{112} = \delta^1_1 \delta^1_1 \delta^2_2 + \epsilon_{112} \epsilon^{112} = 1 \cdot 1 \cdot 1 + 0 \cdot 0 = 1$$

$$T^{213}_{123} = \delta^2_1 \delta^1_2 \delta^3_3 + \epsilon_{213} \epsilon^{213} = 0 \cdot 0 \cdot 1 + (-1) \cdot 1 = -1$$

Since $\mathcal{T}^3_3$ over a vector space of dimension 3 has $3^6 = 729$ components, we typically don't write them all out. Instead, we represent it symbolically or provide a generating rule (as above). However, if we fix some indices, we can extract sub-tensors. For example, fixing $i = 1, l = 1$ yields T^{1jk}_{1mn} which has a smaller number of components, i.e., 3^4.

We can further analyze the tensor space defined above by evaluating an input set of vectors under the assumption $V = \mathbb{R}^3$. To that end, take vectors $u_1, u_2, u_3 \in \mathbb{R}^3$ and $v^1, v^2, v^3 \in (\mathbb{R}^3)^*$. We will show that

$$T(v^1, v^2, v^3, u_1, u_2, u_3) = (u_1 v^1)(u_2 v^2)(u_3 v^3) + det(v^1, v^2, v^3) \cdot det(u_1, u_2, u_3)$$

First, let's work on the $(u_1 v^1)(u_2 v^2)(u_3 v^3)$ term. Using Einstein notation, note that $u_i v^i = u^j_i v^i_j$, for each value of $i = 1,2,3$ and the summation over j (we use this in the derivation below).

In general, we have that $\delta_q^p x_p = x_q$. So, when $\delta_i^l \delta_j^m \delta_k^n$ is applied to our input vectors $(v_l^1 v_m^2 v_n^3)(u_1^i u_2^j u_3^k) = (u_1^i v_l^1)(u_2^j v_m^2)(u_3^k v_n^3)$, we get

$$\delta_i^l \delta_j^m \delta_k^n (u_1^i v_l^1)(u_2^j v_m^2)(u_3^k v_n^3)$$

$$= \left(\delta_i^l u_1^i v_l^1\right)\left(\delta_j^m u_2^j v_m^2\right)\left(\delta_k^n u_3^k v_n^3\right)$$

$$= \left(u_1^l v_l^1\right)(u_2^m v_m^2)(u_3^n v_n^3)$$

$$= (u_1 v^1)(u_2 v^2)(u_3 v^3)$$

Next, let's work on $det(u_1, u_2, u_3) \cdot det(v^1, v^2, v^3)$. From Theorem 78, we have that

$$\epsilon_{ijk} u_1^i u_2^j u_3^k = det \begin{pmatrix} u_1^1 & u_1^2 & u_1^3 \\ u_2^1 & u_2^2 & u_2^3 \\ u_3^1 & u_3^2 & u_3^3 \end{pmatrix} = det(u_1, u_2, u_3)$$

$$\epsilon^{lmn} v_l^1 v_m^2 v_n^3 = det \begin{pmatrix} v_1^1 & v_1^2 & v_1^3 \\ v_2^1 & v_2^2 & v_2^3 \\ v_3^1 & v_3^2 & v_3^3 \end{pmatrix} = det(v^1, v^2, v^3)$$

In summary,

$$T(v^1, v^2, v^3, u_1, u_2, u_3) = (v_l^1 v_m^2 v_n^3)(u_1^i u_2^j u_3^k) T_{lmn}^{ijk}$$

$$= (u_1 v^1)(u_2 v^2)(u_3 v^3) + det(v^1, v^2, v^3) \cdot det(u_1, u_2, u_3)$$

The output scalar depends on the input vectors, but the tensor components T_{lmn}^{ijk} are fixed once bases are chosen for V and V^*. This is true in general for tensors.

Theorem 79. The set $\mathcal{T}_r^s$ of tensors of type (r, s) over vector space V (of dimension n) forms a vector space of dimension n^{r+s}.

Proof: First, we define the set of tensors that will serve as a basis for $\mathcal{T}_r^s$.

For each set of indices $\{i_1, \dots, i_r, k_1, \dots, k_s\}$ such that each index is between 1 and n (the dimension of V), define the tensor $S_{k_1 \dots k_s}^{i_1 \dots i_r} \in \mathcal{T}_r^s$ such that its values on the bases $\{e_i\}$ for V and $\{\omega^k\}$ for V^* are given by

$$S_{k_1 \dots k_s}^{i_1 \dots i_r}\left(\omega^{l_1}, \dots, \omega^{l_s}, e_{j_1}, \dots, e_{j_r}\right) = \delta_{k_1}^{l_1} \dots \delta_{k_s}^{l_s} \delta_{j_1}^{i_1} \dots \delta_{j_r}^{i_r}$$

Let $\{v^1, \dots, v^s, u_1, \dots, u_r\}$ be an arbitrary collection of r vectors in V and s vectors in V^*. Using Einstein notation, we can write $u_i = e_j a_i^j$ where $a_i^j \in \mathbb{R}$ is the j^{th} component of u_i, and we can write $v^k = \omega^l b_l^k$ where $b_l^k \in \mathbb{R}$ is the l^{th} component of v^k.

Using the multilinearity property, we have

$$S^{i_1\ldots i_r}_{k_1\ldots k_s}(v^1,\ldots,v^s,u_1,\ldots,u_r) = S^{i_1\ldots i_r}_{k_1\ldots k_s}\left(\omega^{l_1}b^1_{l_1},\ldots,\omega^{l_s}b^s_{l_s},e_{j_1}a^{j_1}_1,\ldots,e_{j_r}a^{j_r}_r\right)$$

$$= b^1_{l_1}\ldots b^s_{l_s}a^{j_1}_1\ldots a^{j_r}_r S^{i_1\ldots i_r}_{k_1\ldots k_s}\left(\omega^{l_1},\ldots,\omega^{l_s},e_{j_1},\ldots,e_{j_r}\right)$$

$$= b^1_{l_1}\ldots b^s_{l_s}a^{j_1}_1\ldots a^{j_r}_r\delta^{l_1}_{k_1}\ldots\delta^{l_s}_{k_s}\delta^{i_1}_{j_1}\ldots\delta^{i_r}_{j_r}$$

$$= b^1_{k_1}\ldots b^s_{k_s}a^{i_1}_1\ldots a^{i_r}_r$$

To prove that the n^{r+s} tensors $S^{i_1\ldots i_r}_{k_1\ldots k_s}$ form a basis for $\mathcal{T}^s_r$, we need to show that they are linearly independent and span $\mathcal{T}^s_r$.

Regarding linear independence, suppose $c^{k_1\ldots k_s}_{i_1\ldots i_r}S^{i_1\ldots i_r}_{k_1\ldots k_s} = 0$ (i.e., a linear combination of the $S^{i_1\ldots i_r}_{k_1\ldots k_s}$ set equal to zero) where $c^{k_1\ldots k_s}_{i_1\ldots i_r}$ are scalars in $\mathbb{R}$ associated with each multi-index $(i_1,\ldots,i_r,k_1,\ldots,k_s)$. There are n^{r+s} such coefficients, one for each basis tensor.

By the definition of $S^{i_1\ldots i_r}_{k_1\ldots k_s}$, we have that

$$c^{k_1\ldots k_s}_{i_1\ldots i_r}S^{i_1\ldots i_r}_{k_1\ldots k_s}\left(\omega^{l_1},\ldots,\omega^{l_s},e_{j_1},\ldots,e_{j_r}\right)$$

$$= c^{k_1\ldots k_s}_{i_1\ldots i_r}\delta^{l_1}_{k_1}\ldots\delta^{l_s}_{k_s}\delta^{i_1}_{j_1}\ldots\delta^{i_r}_{j_r}$$

$$= c^{k_1\ldots k_s}_{i_1\ldots i_r} = 0$$

So, $c^{k_1\ldots k_s}_{i_1\ldots i_r} = 0$ for every choice of $(l_1,\ldots,l_s,j_1,\ldots,j_r)$. Thus, the only linear combination of the $S^{i_1\ldots i_r}_{k_1\ldots k_s}$ tensors that yields the zero tensor is the trivial one, where all coefficients are zero, and therefore, the collection of tensors $\left\{S^{i_1\ldots i_r}_{k_1\ldots k_s}\right\}$ is linearly independent.

Regarding the span of $\left\{S^{i_1\ldots i_r}_{k_1\ldots k_s}\right\}$, consider the tensor $T^{k_1\ldots k_s}_{i_1\ldots i_r}S^{i_1\ldots i_r}_{k_1\ldots k_s}$ for some arbitrary $T\in\mathcal{T}^s_r$.

The tensors $T^{k_1\ldots k_s}_{i_1\ldots i_r}S^{i_1\ldots i_r}_{k_1\ldots k_s}$ and T yield the same result when applied to any $(r+s)$-tuple $(\omega^{l_1},\ldots,\omega^{l_s},e_{j_1},\ldots,e_{j_r})$, i.e.,

$$T\left(\omega^{l_1},\ldots,\omega^{l_s},e_{j_1},\ldots,e_{j_r}\right) = T^{l_1\ldots l_s}_{j_1\ldots j_r}\text{ (by definition of tensor components)}$$

$$T^{k_1\ldots k_s}_{i_1\ldots i_r}S^{i_1\ldots i_r}_{k_1\ldots k_s}\left(\omega^{l_1},\ldots,\omega^{l_s},e_{j_1},\ldots,e_{j_r}\right) = T^{k_1\ldots k_s}_{i_1\ldots i_r}\delta^{l_1}_{k_1}\ldots\delta^{l_s}_{k_s}\delta^{i_1}_{j_1}\ldots\delta^{i_r}_{j_r} = T^{l_1\ldots l_s}_{j_1\ldots j_r}$$

Thus, they must be equal as multilinear functions over $(V^*\times\ldots\times V^*)\times(V\times\ldots\times V)$. So, any $T\in\mathcal{T}^s_r$ can be expressed as a linear combination $\left\{S^{i_1\ldots i_r}_{k_1\ldots k_s}\right\}$ and therefore, these tensors span $\mathcal{T}^s_r$. ■

As we will describe in Theorem 80, it is possible to write a basis for $\mathcal{T}_r^s$ in terms of the bases $\{e_i\}$ for V and $\{\omega^j\}$ for V^*. But first we need to define something known as the **tensor product**. Take $S \in \mathcal{T}_{r_1}^{s_1}$ and $T \in \mathcal{T}_{r_2}^{s_2}$. In addition, let $u_1, \dots, u_{r_1}, v_1, \dots, v_{r_2} \in V$ and $\alpha^1, \dots, \alpha^{s_1}, \beta^1, \dots, \beta^{s_2}$ be vectors in V^*. The following product is linear in each of its $r_1 + s_1 + r_2 + s_2$ variables

$$S\left(\alpha^1, \dots, \alpha^{s_1}, u_1, \dots, u_{r_1}\right) T(\beta^1, \dots, \beta^{s_2}, v_1, \dots, v_{r_2})$$

We define the tensor product $S \otimes T \in \mathcal{T}_{r_1+r_2}^{s_1+s_2}$ by

$$S \otimes T\left(\alpha^1, \dots, \alpha^{s_1}, \beta^1, \dots, \beta^{s_2}, u_1, \dots, u_{r_1}, v_1, \dots, v_{r_2}\right)$$
$$= S\left(\alpha^1, \dots, \alpha^{s_1}, u_1, \dots, u_{r_1}\right) T(\beta^1, \dots, \beta^{s_2}, v_1, \dots, v_{r_2})$$

The following associative and distributive laws hold for the tensor product:

- $(R \otimes S) \otimes T = R \otimes (S \otimes T)$
- $R \otimes (S + T) = (R \otimes S) + (R \otimes T)$
- $(R + S) \otimes T = (R \otimes T) + (S \otimes T)$
- $(aS) \otimes T = S \otimes (aT) = a(S \otimes T), a \in \mathbb{R}$

However, the tensor product is not commutative.

[**Author's remark**: The symbol $\otimes$ is used (in general, not just this book) for both the Kronecker product and the tensor product. In the event that the two cannot be distinguished by context, we will use $\boxtimes$ for the Kronecker product.]

Making use of the tensor product, we have the following theorem concerning the representation of a basis for $\mathcal{T}_r^s$ in terms of the basis vectors for V and V^*.

Theorem 80. If V has basis $\{\boldsymbol{e_1}, \dots, \boldsymbol{e_n}\}$ and V^* has corresponding basis $\{\boldsymbol{\omega^1}, \dots, \boldsymbol{\omega^n}\}$, then a basis for $\mathcal{T}_r^s$ is given by the collection

$$\left\{\boldsymbol{w^{j_1}} \otimes \dots \otimes \boldsymbol{w^{j_s}} \otimes \boldsymbol{e_{i_1}} \otimes \dots \otimes \boldsymbol{e_{i_r}}\right\}$$

where $\boldsymbol{1 \le j_1, \dots, j_r, i_1, \dots, i_s \le n}$, and thus, $\boldsymbol{dim\, \mathcal{T}_r^s = n^{r+s}}$.

Proof: The discussion preceding Theorem 11.1 in Broida [41] provides a sketch of a proof. There is also a proof in Lee [43], see Corollary 12.12. ■

. . .

In what follows, we derive the change of basis formula for a tensor.

Let T be an (r, s) type tensor over a vector space V where

- $\{e_i\}$ is the original basis for V
- $\{e^i\}$ is the corresponding dual basis for V^* such that $e^i e_j = \delta_j^i$
- Let $\{\hat{e}_i\}$ and $\{\hat{e}^j\}$ be new bases for V and V^* respectively.

The change of bases is given by the following formula:

$$\hat{e}_i = A_i^j e_j, \qquad \hat{e}^i = (A^{-1})_j^i e^j$$

Note that the covariant (i.e., dual) indices transform with the inverse of the change of basis matrix.

The components of T with respect to the original bases $\{e_i\}, \{e^i\}$ are

$$T^{i_1 \dots i_s}_{j_1 \dots j_r} = T(e_{j_1}, \dots, e_{j_s}, e^{i_1}, \dots, e^{i_r})$$

Next, we evaluate the components of T in the new basis, i.e.,

$$T^{a_1 \dots a_s}_{b_1 \dots b_r} = T(\hat{e}_{b_1}, \dots, \hat{e}_{b_s}, \hat{e}^{a_1}, \dots, \hat{e}^{a_r})$$

$$= T\left(A^{j_1}_{b_1} e_{j_1}, \dots, A^{j_s}_{b_s} e_{j_s}, (A^{-1})^{a_1}_{i_1} e^{i_1}, \dots, (A^{-1})^{a_r}_{i_r} e^{i_r}\right)$$

$$= A^{j_1}_{b_1} \cdots A^{j_s}_{b_s} (A^{-1})^{a_1}_{i_1} \cdots (A^{-1})^{a_r}_{i_r} \cdot T\left(e_{j_1}, \dots, e_{j_s}, e^{i_1}, \dots, e^{i_r}\right)$$

$$= A^{j_1}_{b_1} \cdots A^{j_s}_{b_s} (A^{-1})^{a_1}_{i_1} \cdots (A^{-1})^{a_r}_{i_r} \cdot T^{i_1 \dots i_s}_{j_1 \dots j_r}$$

A multidimensional array needs to follow the above transformation law regarding a change of basis, or it is not considered to be a tensor. This is the defining property of a tensor: its components must transform in this way under a change of basis.

6.2 Tensors as Multidimensional Arrays

An alternative yet equivalent definition of a tensor can be formulated using multidimensional arrays, see the Wikipedia article “Tensor” [40]. A multidimensional array becomes a tensor when its components satisfy a consistent transformation law under a change of basis of the underlying vector space.

Just as a vector in an n-dimensional space is represented by a one-dimensional array of n components relative to a chosen basis, a tensor of any order is represented by a multidimensional array with respect to a basis. For instance, a type $(1,1)$ tensor is represented as a two-dimensional $n \times n$ array with respect to given basis.

The numbers in the multidimensional array are known as the components of the tensor. Each component is labeled by a set of indices that indicate its position within the array. These indices are typically written as subscripts and/or superscripts following the symbolic name of the tensor. For example, the components of a second-order tensor T could be denoted T_{ij}, where i and j are indices running from 1 to n, or alternatively as T^i_j. Whether an index is displayed as a superscript or subscript depends on its transformation properties, discussed later.

While T_{ij} and T^i_j can both be expressed as $n \times n$ matrices, and may even share numerical values with respect to a fixed basis, they obey different transformation laws under a change of basis. As such, they represent different mathematical objects. Consequently, adding them together would be inappropriate, despite the superficial similarity of their array structures.

The total number of indices required to uniquely identify each component of a tensor is called the rank (or order) of the tensor. This should not be confused with the dimension of the underlying vector space. A tensor of rank r can be viewed as an r-way array, which is why tensors are sometimes described as multidimensional arrays or multiway arrays.

Example: The following is a Python program that multiplies two $3 \times 3 \times 3$ tensors. It makes use of the Python tensorflow package. The operation for matrix multiplication is the @ sign. The two tensors in the program are named $t1$ and $t2$. Their product is $t3$.

```
import tensorflow as tf
# Tensor (3D Tensor) - Higher dimensional data
t1 = tf.constant(
[
[[1,2,1],[3,3,4],[-4,8,5]],
[[5,2,6],[7,8,8],[1,1,2]],
[[3,-2,4],[1,-7,6],[3,3,3]]
]
)
t2 = tf.constant(
[
[[3,2,-1],[1,3,4],[1,-3,2]],
[[-1,2,3],[1,8,1],[1,3,2]],
[[1,5,-1],[4,-7,1],[3,2,3]]
]
)
print(t1,t2)
t3=t1@t2
print(t3)
```

The output of the program is as follows. Three matrix multiplications are required to compute the product of $t1$ and $t2$, i.e., multiply the first matrix in $t1$ times the first matrix in $t2$, multiply the second matrix in $t1$ times the second matrix in $t2$ and then multiply the third matrix in $t1$ times the third matrix in $t2$.

```
Tensor t1 tf.Tensor(
[[[ 1  2  1]
  [ 3  3  4]
  [-4  8  5]]

 [[ 5  2  6]
  [ 7  8  8]
  [ 1  1  2]]

 [[ 3 -2  4]
  [ 1 -7  6]
  [ 3  3  3]]], shape=(3, 3, 3), dtype=int32)
```

Tensor t2 tf.Tensor(
[[[3 2 -1]
  [1 3 4]
  [1 -3 2]]

 [[-1 2 3]
  [1 8 1]
  [1 3 2]]

 [[1 5 -1]
  [4 -7 1]
  [3 2 3]]], shape=(3, 3, 3), dtype=int32)

Tensor t3= t1@t2 tf.Tensor(
[[[6   5   9]
  [16   3 17]
  [1   1 46]]

 [[3 44 29]
  [9 102 45]
  [2 16   8]]

 [[7 37   7]
  [-9 66 10]
  [24   0   9]]], shape=(3, 3, 3), dtype=int32)

. . .

Just as the components of a vector change under a change of basis in a vector space, the components of a tensor also transform according to a rule determined by the type (or variance) of the tensor. Each index in a tensor transforms independently, either covariantly or contravariantly.

To illustrate this, consider a change of basis in a vector space V where the new basis vectors $\hat{e}_i$ are expressed in terms of the old basis vectors e_j as follows (using Einstein notation):

$$\hat{e}_i = A_i^j e_j$$

Here, A_i^j are the entries in the change of basis matrix A, with the index j referring to the original basis and i to the new basis. This follows the covariant transformation law, because the basis vectors are transformed by the change-of-basis matrix itself.

For the dual space V^*, the basis vectors e^j transform with the inverse of the matrix A, i.e.,

$$\hat{e}^i = (A^{-1})_j^i e^j$$

The above follows the contravariant transformation law, because the basis vectors transform by the inverse of the change of basis matrix.

More generally, the components of a tensor with multiple indices transform according to a combination of contravariant and covariant rules – one for each index. If an index transforms with the inverse of the change-of-basis matrix, it is said to be **contravariant**, and is written as a superscript. If it transforms with the matrix itself, it is **covariant**, and is written as a subscript.

The main distinction between defining a tensor as a multilinear map versus a multidimensional array lies in how the change-of-basis formula is treated. In the multilinear map definition, the change-of-basis formula can be derived (as shown earlier), whereas in the multidimensional array definition, it must be explicitly included as part of the definition.

The elements of a tensor transform in a specific manner when the basis changes, guaranteeing that the tensor itself remains unchanged (independent of coordinates). This transformation rule is what sets tensors apart from generic multidimensional arrays – **not every array qualifies as a tensor**, because tensors must adhere to precise transformation laws.

In the context of Machine Learning (ML) and numerical computing, the term "tensor" is often used informally to refer to multi-dimensional data structures that generalize scalars, vectors, and matrices. These structures include:

- Vectors (1D arrays): Such as feature vectors in a model.
- Matrices (2D arrays): Used for representations like images (height × width) or weight matrices in neural networks.
- Higher-dimensional arrays: For example, a mini-batch of RGB images (batch × height × width × channels) or input tensors in transformer models.

However, these are not tensors in the strict mathematical sense as defined in differential geometry or physics. In those fields, a tensor is not just a container for numbers—it is an object that follows specific transformation rules under changes of coordinate systems. A true tensor's components must adjust in a precise way when the basis (or reference frame) changes, preserving its geometric or physical meaning.

By contrast, in ML, tensors are simply efficient numerical containers with no inherent transformation laws. They serve as a convenient way to store and manipulate multi-dimensional data, but they do not necessarily respect the rigorous mathematical properties of tensors in physics or differential geometry.

This relaxed usage is practical for ML, where the focus is on computation rather than coordinate-invariant properties. However, it’s important to recognize the distinction when moving between fields like deep learning and theoretical physics.

7 Non-Negative Matrices

Action expresses priorities. – Mahatma Gandhi

7.1 Overview

In this section, we will present results related to non-negative matrices. A real matrix A whose element (i, j) is denoted a_{ij} is a **non-negative matrix** if $a_{ij} \geq 0$ for every element of A. In this case, we write $A \geq 0$. If $a_{ij} > 0$ for every element of a real matrix A, then A is said to be a **positive matrix**, and we write $A > 0$. We can extend these definitions to row and column vectors, since they are matrices.

For matrices A and B of the same dimensions, we can have expressions such as $A \geq B$ if and only if $(A - B) \geq 0$.

A square non-negative matrix A is said to be **primitive** if there exists a positive integer n such that $A^n > 0$. If B is of the same dimensions as a primitive matrix A, and B has positive elements and zeros in the same positions as A, then B is also primitive.

An **incidence matrix** $\tilde{A}$ for a non-negative matrix A is formed by replacing all the positive elements of A by ones. In the case of square matrices, the matrix $\tilde{A}$ is primitive if and only if A is primitive.

Non-negative matrices are widely used in various fields due to their interpretability and applicability to real-world data that is naturally non-negative. Here are some key applications:

1. Data Analysis & Machine Learning
    a. Non-negative Matrix Factorization (NMF): Used for dimensionality reduction, feature extraction, and topic modeling (e.g., in text mining, where documents are represented as non-negative word counts).
    b. Recommender Systems: User-item interaction matrices (e.g., ratings, clicks) are non-negative, and NMF or collaborative filtering techniques leverage this structure.
    c. Image Processing: Pixel intensities are non-negative, making NMF useful for parts-based representation (e.g., facial recognition).
2. Statistics & Probability
    a. Markov Chains: Transition matrices (where entries represent probabilities) are non-negative.
    b. Stochastic Processes: Covariance matrices and other statistical measures often involve non-negative values.
3. Network Science and Graph Theory
    a. Adjacency Matrices: Represent connections in graphs (e.g., social networks, web links) where weights are non-negative.
    b. Page Rank Algorithms: Used by search engines to rank web pages based on non-negative link matrices.

4. Economics and Operations Research
    a. Input-Output Models (Leontief Models): Economic systems where production relationships are represented as non-negative matrices.
    b. Transportation & Flow Networks: Matrices describe capacities, distances, or flows between nodes.
5. Biology and Bioinformatics
    a. Gene Expression Analysis: Microarray data is non-negative, and NMF is used to identify patterns or clusters.
    b. Ecological Modeling: Species interaction matrices (e.g., predator-prey relationships) use non-negative entries.
6. Computer Vision
    a. Hyperspectral Imaging: Unmixing spectral data into constituent materials using NMF.
    b. Object Detection: Non-negative constraints improve interpretability in feature learning.
7. Signal Processing
    a. Source Separation: Separating audio or mixed signals (e.g., music, speech) into additive components (e.g., NMF in spectrograms).
    b. Sparse Coding: Representing signals with non-negative bases.

7.2 Perron-Frobenius Theorem for Primitive Matrices

A key (or perhaps "the key") theorem concerning non-negative matrices is the Perron-Frobenius theorem.

Theorem 81. (Perron-Frobenius Theorem for Primitive Matrices) If A is an $n \times n$ real primitive matrix, then there exists an eigenvalue r of A such that:

(1) $r > 0, r \in \mathbb{R}$

(2) Associated with r are strictly positive left and right eigenvectors.

(3) $r > |\lambda|$ ***for any eigenvalue*** $\lambda \neq r$

(4) The eigenvectors associated with r are unique to constant multiples.

(5) If $0 \leq B \leq A$ ***and*** β ***is an eigenvalue of*** B***, then*** $|\beta| \leq r$***. Further,*** $|\beta| = r$ ***implies*** $B = A$***.***

(6) r ***is a simple root of the characteristic equation for*** A***.***

Proof: The Wikipedia article "Perron-Frobenius theorem" [44] discussed several approaches for the proof of this theorem. We will focus our effort on understanding each part of the theorem. ■

Concerning the theorem:

(1) Real non-negative primitive matrices can possibly have complex eigenvalues, but the theorem tells us that there is at least one positive real eigenvalue.

(2) In linear algebra (as presented earlier in this book and in general), one typically uses right eigenvectors, e.g., $Ax = rx$. A left eigenvector would satisfy $yA = ry$. The theorem tells us that A has strictly positive left and right eigenvectors associated with eigenvalue r.

(3) This part of the theorem tells as that r is larger than the absolute value or modulus of any other eigenvalue (including possible complex eigenvalues). Recall that the modulus of complex number $z = a + ib$ is $|z| = \sqrt{a^2 + b^2}$. The **spectral radius** of a square matrix is the maximum of the absolute values or moduli of its eigenvalues. Thus, r is the spectral radius of A.

(4) This item is effectively telling us that eigenspace associated with r is of dimension 1, i.e., r has geometric multiplicity 1. [Recall from Section 2.14 that algebraic multiplicity refers to the number of times an eigenvalue appears as a root of a matrix's characteristic polynomial, while geometric multiplicity is the number of linearly independent eigenvectors associated with that eigenvalue. The geometric multiplicity is always less than or equal to the algebraic multiplicity.]

(5) It is not necessary for B to also be primitive.

(6) While Item (4) tells us that r has geometric multiplicity 1, this item tells us that r has algebraic multiplicity 1.

Many other theorems follow from the Perron-Frobenius theorem. We list a few here.

Let $\mathbf{1}$ be a column vector all of whose entries are 1. Einstein notation is used in the statement of the following theorem.

Theorem 82. If A is a primitive matrix, then $\min_i A_{ij}\mathbf{1}_j \leq r \leq \max_i A_{ij}\mathbf{1}_j$ where r is the eigenvalue of A of maximum value. Similarly, $\min_j \mathbf{1}_i A_{ij} \leq r \leq \max_j \mathbf{1}_i A_{ij}$.

Proof: See Chapter 1, Corollary 1 of Seneta [45]. ■

In words, the above theorem says that r is between the minimum and maximum row sum of A, and also between the minimum and maximum column sum of A.

Theorem 83. If v^T and w are the positive left and right eigenvectors corresponding the maximum value eigenvalue r of primitive matrix A, normalized such that $v^T w = \mathbf{1}$, then

$$\frac{\mathbf{adj}(rI - A)}{\frac{d}{dr}[\det(rI - A)]} = wv^T$$

Proof: See Chapter 2, Corollary 1 of Seneta [45]. ■

7.3 Perron-Frobenius Theorem for Irreducible Matrices

The Perron-Frobenius theorem can be applied under a weaker condition known as matrix irreducibility. Before we get to the modified version of the theorem, we need to define several preliminary concepts.

The **directed graph** (or digraph) associated with an $n \times n$ matrix $A = [a_{ij}]$ is a graph that represents the non-zero structure of the matrix. In particular, the graph has n vertices, labeled $1,2,\ldots,n$, where each vertex corresponds to a row (or column) of the matrix. In addition, there is a directed edge from vertex i to vertex j if and only if the matrix entry $a_{ij} \neq 0$. The edge is directed from i to j, indicating a relationship or influence from row i to column j. In some contexts, edges may be assigned weights equal to the values of a_{ij}, but for the purpose of determining irreducibility, only the presence of non-zero entries matters.

For example, consider the matrix:

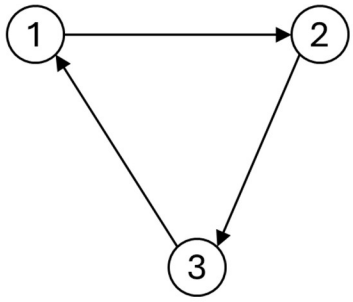
$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Vertices: $\{1,2,3\}$

Edges:

- $a_{12} = 1 \neq 0$, and so, there is an edge $1 \to 2$
- $a_{23} = 1 \neq 0$, and so, there is an edge $2 \to 3$
- $a_{31} = 1 \neq 0$, and so, there is an edge $3 \to 1$

Graph: The digraph forms a cycle $1 \to 2 \to 3 \to 1$, as shown below.



A directed graph is **strongly connected** if, for every pair of distinct vertices $\{i, j\}$, there exists a directed path from vertex i to j, and from vertex j to i. A **directed path** is a sequence of edges where the endpoint of one edge is the starting point of the next, respecting the direction of the edges. Formally, a digraph with vertices $\{1,2, \dots, n\}$ is strongly connected if, for any $i \neq j$, there is a sequence of vertices $i = v_0, v_1, \dots, v_k = j$ such that there are edges $v_0 \to v_1, v_1 \to v_2, \dots, v_{k-1} \to v_k$, and similar, concerning a path from j to i. Strong connectivity means the graph is "fully reachable", i.e., you can travel from any vertex to any other vertex following the directed edges.

The matrix A in the example above is strongly connected since it has a cycle transversing all the vertices. The matrix B below is not strongly connected since, for example, there is no path from v_2 to v_3.

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\dots$$

An **irreducible matrix** is a square matrix that cannot be permuted into a block upper triangular form by any permutation of its rows and columns. This property ensures that the matrix cannot be decomposed into smaller, independent submatrices.

More precisely, an $n \times n$ matrix $A = [a_{ij}]$ is irreducible if, for any permutation matrix P, the matrix $P^{-1}AP$ cannot be written in the block matrix form

$$P^{-1}AP = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

where B and D are square submatrices, and 0 is a zero matrix.

Equivalently, a square matrix A is irreducible if and only if its associated directed graph is strongly connected. This is because:

- If the graph is strongly connected, there is no way to partition the vertices into two sets where there are no edges from one set to the other, preventing the matrix from being permuted into block upper triangular form.
- Conversely, if the graph is not strongly connected, the vertices can be partitioned into subsets (e.g., using strongly connected components), and the matrix can be permuted into a block form, making it reducible.

We can now state a version of the Perron-Frobenius theorem for irreducible matrices.

Theorem 84. If A is an $n \times x$ non-negative irreducible real matrix, then the following is true:

(1) Perron Root: A has a real positive eigenvalue r, called the Perron root or spectral radius, such that $r \geq |\lambda|$ for any other eigenvalue λ of A.

(2) Simple Eigenvalue: The Perron root r is a simple eigenvalue, i.e., its algebraic multiplicity is 1.

(3) Positive Eigenvector: There exists a positive eigenvector v, i.e., $v > 0$, corresponding to r and this eigenvector is unique up to a positive scalar multiple.

(4) Dominance: No other eigenvalue of A has an eigenvector with all positive components.

(5) Spectral Radius: The Perron root r satisfies $r = \rho(A)$, where $\rho(A) = max\{|\lambda|: \lambda$ is an eigenvalue of $A\}$.

(6) ***Periodicity (for primitive matrices): If A is primitive (a stronger condition than irreducibility), then r is the only eigenvalue of maximum modulus.***

Proof: See Theorem 1.5 in Seneta [45]. ■

Concerning the Perron root, note that in the above theorem, we have $r \geq |\lambda|$ but in Theorem 81 (3), we have $r > |\lambda|$. Other than this difference, the two theorems yield the same properties.

8 Markov Chains

Knowledge is having the right answer. Intelligence is asking the right question.

Stephen Hawking.

8.1 Definition

A **Markov chain** is a stochastic (or random) process describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. A countably infinite sequence, in which the process changes state at discrete time steps, defines a discrete-time Markov chain. A continuous-time process is called a continuous-time Markov chain. Our focus here will be on discrete-time Markov chains.

More formally, let $\{X_n\}$ for $n = 0, 1, 2, \ldots$ be a sequence of random variables with a finite or countably infinite state space $I = \{i_0, i_1, i_2, \ldots\}$. The random variable X_n represents the state of a random process system at time n. A random process is a Markov Chain if for each $n = 0, 1, 2, \ldots$

$$P(X_{n+1} = i_{n+1} \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\} = P(X_{n+1} = i_{n+1} \mid X_n = i_n\}$$

for all possible values of $i_0, i_1, i_2 \ldots \in I$. In other words, given the past history of a process, the transition to the next step only depends on the present state. This is known as the Markov property [46].

If the transition probability $P(X_{n+1} = j \mid X_n = i\} = p_{ij}$ does not depend on n (as emphasized in the shorthand notation p_{ij}), then the Markov chain is said to be **time-homogeneous**.

Let $S = \{i_0, i_1, i_2, \ldots\}$ be the set of states for a given Markov chain. This is known as the **state space** for the Markov chain. If a Markov chain is in a given state (say i), then the sum of the transition probabilities from i must add to 1, i.e.,

$$\sum_{j \in S} p_{ij} = 1$$

Two states in a Markov chain are said to **communicate** if they are reachable from one another (bidirectional) by a sequence of transitions with positive probability. This is an equivalence relation [47] which yields a set of communication classes. A Markov chain is said to be **irreducible** if there is one communication class, i.e., the entire state space.

Consider a Markov chain $\{X_n\}$ with state space $S = \{i_1, i_2, i_3, \ldots\}$ and transition probabilities given by $P(X_{k+1} = j \mid X_k = i\} = p_{ij}$. The transition probabilities can be represented as the matrix T below:

$$T = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots \\ p_{21} & p_{22} & p_{23} & \cdots \\ p_{31} & p_{32} & p_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

If the number of states is n, then the above matrix would be $n \times n$.

T is known as the transition matrix for the Markov chain. Each row represents the probabilities of transitioning from a given state to any of the other states. For example, row 2 has the probabilities of transitioning from state i_2 to each of the other states. As defined, each row is a discrete probability distribution. A matrix with all non-negative entries and whose rows sum to 1 is said to be **row-stochastic**.

It is possible to compute the probabilities of going from one state to another in a n steps, i.e.,

$$p_{ij}^{(n)} = P(X_{k+n} = j \mid X_k = i \}$$

In fact, $p_{ij}^{(n)}$ is the (i, j) element of T^n (see the section concerning the Chapman-Kolmogorov Equations in [48]). To be clear, $p_{ij}^{(n)}$ is the probability of going from state i at step k and being in state j at step n. This includes the possibility of earlier visits to state j. This is different from the probability of the first visit to state j starting from state i.

The probabilities in transition matrix T are all conditioned on a given state. It is also possible to compute the probability of being in a given state after n steps based on an initial probability distribution. Let $p_j^{(0)}$ be the probability that a given Markov chain is initially in state j. We can write the set of initial probabilities in the form of a row vector, i.e.,

$$p^{(0)} = (p_1^{(0)}, p_2^{(0)}, p_3^{(0)}, \dots)$$

The probability of being in state j at step n given initial probability distribution $p^{(0)}$ is represented by the notation $p_j^{(n)}$. Using the law of total probability, we have

$$p_j^{(n)} = \sum_{i=1}^{\infty} p_i^{(0)} p_{ij}^{(n)}$$

The above formula is $p^{(0)}$ times column j in T^n. This is true for all values of j. Thus, we have

$$p^{(n)} = p^{(0)} T^n$$

where we define $p^{(n)}$ as follows

$$p^{(n)} = (p_1^{(n)}, p_2^{(n)}, p_3^{(n)}, \dots)$$

$p_j^{(n)}$ is known as absolute or unconditional probability of being in state j at step n, given the initial distribution $p^{(0)}$.

We can also derive a formula for $p^{(n)}$ based on $p^{(n-1)}$ and T. From the law of total probability, we have

$$p_j^{(n)} = \sum_{i=1}^{\infty} p_i^{(n-1)} p_{ij}^{(1)}$$

In matrix notation, the above can be written as

$$p^{(n)} = p^{(n-1)} T$$

8.2 Examples

8.2.1 Bernoulli Scheme

Consider a sequence of independent trials each with fixed probability p of occurring (i.e., being successful) and therefore, a probability $q = 1 - p$ of not occurring (i.e., of failing). We equate success with the number 1 and failure with the number 0. In this way, our state space is $\{0,1\}$. This gives us a two-state, time-homogeneous Markov chain. At time k (i.e., at trial k), we can represent the possible transitions via the following matrix

$$T = \begin{bmatrix} q & p \\ q & p \end{bmatrix}$$

If the previous trial (i.e., $k - 1$) was a failure (state 0), the first row of T tells us that the probability of failure in trial k is q and the probability of success is p. If the previous trial was a success, the second row of T tells us the probability of failure in trial k is q and the probability of success is p. The rows of the transition matrix are identical because the trials are independent, making the state transitions independent of the current state.

8.2.2 Random Walk Between Two Barriers

An entity (e.g., a person) starts at any of the points $\{0,1,2, \dots, s\}$ on a ruled line. Each move (step) of the entity constitutes a transition. The transitions follow these rules:

- If the entity reaches point 0 it remains there with probability a, and moves to state 1 on the next step with probability $1 - a$.
- If the entity reaches point s it remains there with probability b, and moves to state $s - 1$ on the next step with probability $1 - b$.
- For points other than 0 or s, the entity moves to a lower number state with probability q on the next step and to a higher number state with probability $p = 1 - q$.

So described, we have a homogeneous Markov chain with finite state space $\{0,1,2, \dots, s\}$. The transitions can be represented by the following matrix:

$$T = \begin{bmatrix} a & 1-a & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & q & 0 & p & \dots & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1-b & b \end{bmatrix}$$

In the above, row 1 represents the transition probabilities from state 0 to the other states, row 2 represents the transition probabilities from state 1 to the other states, and so on.

8.2.3 Markov Chain with an Infinite State Space

Consider a system that either moves to a higher state or goes back to state 0 after a transition. For example, this could model the operating system on a personal computer that is turned “On” each morning, where it could continue for another day, or crash and need to be rebooted. The system has a variable lifetime, whose length (measured in discrete units) has probability distribution $\{p_i\}, i = 1,2,3, \dots$

The state transitions are described by the following transition matrix.

$$\begin{bmatrix} p_1 & 1-p_1 & 0 & 0 & 0 & \dots \\ \frac{p_2}{1-p_1} & 0 & \frac{1-p_1-p_2}{1-p_1} & 0 & 0 & \dots \\ \frac{p_3}{1-p_1-p_2} & 0 & 0 & \frac{1-p_1-p_2-p_3}{1-p_1-p_2} & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

The state space is $\{0,1,2,\dots\}$. In the above matrix, row i gives the transition probabilities from state i to the other states (assuming we start the row numbering from 0).

This is a time-homogeneous Markov chain since the transitions do not depend on the time n, i.e., the number of previous transitions.

Practical Interpretation in the Example: For the computer system, starting "On" each morning (state 0), p_1 is the immediate crash probability. If it survives (moves to state 1), the crash probability from there is scaled by the remaining mass of the distribution, reflecting updated odds given that it didn't crash early. This avoids over- or under-estimating failure risks in later stages, making the model realistic for systems with non-memoryless lifetimes (unlike a simple geometric distribution, which would have constant probabilities).

8.2.4 Pólya's Urn Scheme

Assume that we have a white and b black balls in an urn. Let $a + b = N$. At each transition, we remove a ball from the urn at random and take note of its color. Before drawing the next ball, we replace the removed ball along with s more balls of the same color.

If the number of white balls selected in the first r drawings is i, we define the system to be in state i. Suppose we are in state i $(i \leq r)$ after drawing number r. This implies $r - i$ black balls have been drawn so far, the number of white balls in the urn is $a + is$, and the number of black balls in the urn is $b + (r - i)s$. At the next drawing, transition to state $i + 1$ occurs with probability

$$p_{i,i+1}(r+1) = \frac{a+is}{N+rs}$$

The probability of transitioning to state i is

$$p_{i,i}(r+1) = \frac{b+(r-i)s}{N+rs} = 1 - p_{i,i+1}(r+1)$$

The key point of this example is that the transition probabilities depend on the number of trials r (selections from the urn) and thus, we have a non-homogeneous Markov chain.

8.3 Irreducible Markov Chains

An **irreducible Markov chain** is one whose transition matrix is irreducible. This is equivalent to the earlier definition that we provided for an irreducible Markov chain.

Let $\{X_n\}$ be an irreducible Markov chain with transition matrix T and let $\mathbf{1}$ be a column vector with 1 in each position. Since the rows of a transition matrix of a Markov chain always add to 1, we have that $T\mathbf{1} = 1 \cdot \mathbf{1}$. Thus, 1 is an eigenvalue of T, and $\mathbf{1}$ a corresponding eigenvector. Since all the row sums of T are equal, we have (by Theorem 82) that $r = 1$ is the Perron eigenvalue of T, and $\mathbf{1}$ can be taken as the corresponding right eigenvector. Let v be a row vector, normalized such that $v\mathbf{1} = 1$, be the corresponding left eigenvector. We have that $vT = v$ where v is the row vector that represents a probability distribution.

Any initial probability distribution $p^{(0)}$ is said to be **stationary** if $p^{(0)} = p^{(n)}, n = 1,2,3, \ldots$ which is equivalent to $p^{(0)}T = p^{(0)}$.

A Markov chain with such an initial distribution is said to be stationary.

For example, the following transition matrix T (representing a Markov chain) has stationary probability $v = \frac{1}{211}(16,24,36,54,81)$, i.e., $v = vT$.

$$T = \begin{bmatrix} .4 & .6 & 0 & 0 & 0 \\ .4 & 0 & .6 & 0 & 0 \\ 0 & .4 & 0 & .6 & 0 \\ 0 & 0 & .4 & 0 & .6 \\ 0 & 0 & 0 & .4 & .6 \end{bmatrix}$$

Theorem 85. A finite irreducible Markov chain has a unique stationary distribution given by the solution v to $vT = v, v\mathbf{1} = 1$.

Proof: At the of Section 8.1, we derived the formula $p^{(n)} = p^{(n-1)}T, n = 0,1,2, \ldots$

If we let $v = p^{(0)}$ where v is the left eigenvector of the Perron eigenvalue (which we know to exist from the Perron-Frobenius theorem), then

$$p^{(1)} = p^{(0)}T = vT = v = p^{(0)}$$

$$p^{(2)} = p^{(1)}T = p^{(0)}T = p^{(0)}$$

and in general, $p^{(n)} = p^{(0)}, n = 1,2, \ldots$

Conversely, if $p^{(0)}$ is a stationary distribution of the given Markov chain, then

$$p^{(0)} = p^{(0)}T, \qquad p^{(0)} \geq 0, \qquad p^{(0)}\mathbf{1} = 1$$

By the uniqueness of the Perron-Frobenius left eigenvector of T, $p^{(0)} = v$. ■

The following theorem is important in Markov chain theory since it states that for a primitive Markov chain the probability distribution of X_k approaches v as $k \to \infty$, i.e., $\lim_{k\to\infty} T^k = \mathbf{1}v$ where v is the left eigenvector corresponding to the Perron eigenvalue. The result is independent of $p^{(0)}$. Further, the rate of convergence is very fast. Thus, after a relatively few transitions, past history becomes irrelevant, and the chain approaches a stationary state.

Theorem 86. (Ergodic theorem for primitive Markov chains) For a primitive Markov chain with unique stationary distribution v and transition matrix T

$$\lim_{k\to\infty} T^k = \mathbf{1}v$$

The rate of convergence is geometric.

Proof: See Theorem 4.2 in Seneta [45]. ■

Since $\mathbf{1}$ is a column vector and v is a row vector, $\mathbf{1}v$ is a matrix.

Example 1. Consider the irreducible Markov chain with state space $\{0,1,2,\dots\}$ and transition matrix

$$T = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & 0 & 0 & \dots \\ \frac{2}{3} & 0 & \frac{1}{3} & 0 & \dots \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} & \dots \\ 0 & 0 & \frac{2}{3} & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

Next, we solve $vT = v$ to find the stationary probability distribution.

$$v_0 = \frac{2}{3}v_0 + \frac{2}{3}v_1$$

$$v_i = \frac{1}{3}v_{i-1} + \frac{2}{3}v_{i+1}, \qquad i \geq 1$$

This gives us

$$v_1 = \frac{1}{2}v_0$$

$$v_1 = \frac{1}{3}v_0 + \frac{2}{3}v_2 \Rightarrow\ v_2 = \left(\frac{1}{2}\right)^2 v_0$$

$$v_2 = \frac{1}{3}v_1 + \frac{2}{3}v_3 \Rightarrow v_3 = \left(\frac{1}{2}\right)^3 v_0$$

$$\dots$$

$$v_n = \left(\frac{1}{2}\right)^n v_0$$

Thus, $v = v_0(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots)$. We normalize v so that the sum of its entries equals 1, and so, we set $v_0 = \frac{1}{2}$. Thus, $v_n = \left(\frac{1}{2}\right)^{n+1}$.

Example 2. In an example similar to the previous example, we limit the state space to $\{0,1,2,3\}$ and adjust the transition matrix as follows:

$$T = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Solving $vT = v$ and normalizing v, we get $v = (\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8})$. Listed below is T raised to the powers $4,8,16$ and 32. As one can see, the rows are converging to v^T.

0.592593 0.246914 0.111111 0.0493827
0.493827 0.320988 0.0740741 0.111111
0.444444 0.148148 0.209877 0.197531
0.197531 0.222222 0.197531 0.382716

0.532236 0.25301 0.117208 0.0975461
0.50602 0.260631 0.116141 0.117208
0.468831 0.232282 0.143423 0.155464
0.390184 0.234415 0.155464 0.219936

0.504315 0.250695 0.123742 0.121248
0.50139 0.25041 0.124458 0.123742
0.494969 0.248917 0.126667 0.129447
0.484991 0.247484 0.129447 0.138077

0.500082 0.250014 0.124975 0.124928
0.500028 0.250005 0.124992 0.124975
0.499902 0.249983 0.125029 0.125086
0.499713 0.249951 0.125086 0.12525

8.4 Classification of States

Recall that two states (i and j) of a Markov chain are said to communicate if they are reachable from one another by a sequence of transitions with positive probability. This is equivalent to the condition that $p_{i,j}^{(n)} > 0$ for some value of $n \geq 1$. A Markov chain may be divided into collections of states that communicate among themselves. We state this more formally in the following theorem.

Theorem 87. A Markov chain with a finite or countably infinite number of states can be uniquely decomposed into a collection of disjoint sub-chains $C_1, C_2, \ldots$ whose union is the entire Markov chain such that all states within a sub-chain communicate with each other but with no other states.

Each sub-chain is known as a **communication class** of the Markov chain.

A Markov chain with only one communication class is referred to as an **irreducible Markov chain**.

For the gambler's ruin Markov chain [50], we have three communications classes, i.e., $\{0\}, \{1,2, \ldots, a-1\}$ and $\{a\}$.

For the unbounded, 1-dimensional random walk, there is but one communication class and so the Markov chain is irreducible.

The Markov chain in Figure 6 has three communication classes, i.e., $\{0,1,2\}, \{3\}$ and $\{4,5,6\}$.

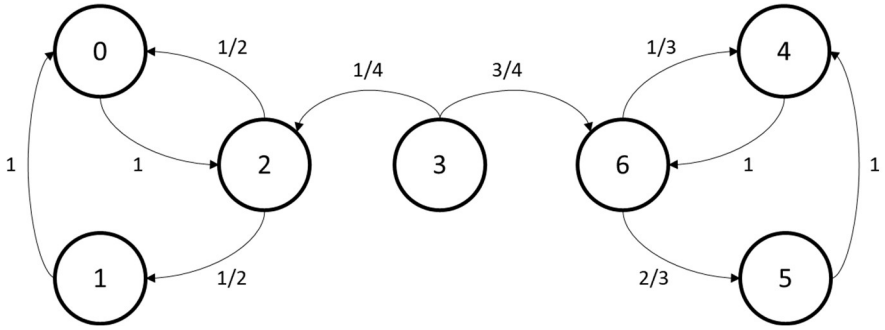


Figure 6. Markov chain with three communication classes

. . .

In addition to the probability of a Markov chain returning to a given state, we are also interested in the number of steps it takes to return. To that end, we define the return time to state i as follows:

$$\tau_{i,i} = \min \{n \geq 1 \text{ such that } X_n = i \text{ given that } X_0 = i\}$$

If no such n exists (i.e., the Markov chain never returns to state i), then $\tau_{i,i} = \infty$.

Further, we define $f_{i,i} = P(\tau_{i,i} < \infty)$ as the probability of ever returning to state i given that the chain started in state i.

A state i is defined to be **recurrent** if $f_{i,i} = 1$; and **transient** if $f_{i,i} < 1$. [Some sources use the term "persistent" to mean the same thing as "recurrent."]

By the Markov property, each time a Markov chain visits a state i it will return with the same probability $f_{i,i}$. So, if $f_{i,i} = 1$, state i will be visited an infinite number of times (thus, the term "recurrent"). If $f_{i,i} < 1$, state i will only be visited a finite number of times.

The total number of visits to state i, starting from state i, is given by the equation

$$N_{i,i} = \sum_{n=0}^{\infty} I\{X_n = i | X_0 = i\}$$

[The indicator function $I\{A\}$ equals 1 if A occurs and 0 if A does not occur. So, $P(I\{A\} = 1) = P(A)$ and $P(I\{A\} = 0) = P(\neg A)$. Further,

$$E(I\{A\}) = 1 \cdot P(A) + 0 \cdot P(\neg A) = P(A)$$

So, there is a one-to-one correspondence between the expected value of the indicator function of an event and the probability of that event.]

The random variable $N_{i,i}$ has a geometric distribution when $f_{i,i} < 1$, i.e.,

$$P(N_{i,i} = n) = f_{i,i}^{n-1}(1 - f_{i,i}), \qquad n \geq 1$$

The above equation can be interpreted as the probability of starting in state i, returning $n - 1$ times, and never returning, for a total of n visits (including the initial visit).

The expected number of visits to state i is $E(N_{i,i}) = \frac{1}{1-f_{i,i}}$ (a known result for geometric distributions). This gives us an equivalent definition for recurrent and transient states:

- State i is recurrent ($f_{i,i} = 1$) if and only if $E(N_{i,i}) = \infty$.
- State i is transient ($f_{i,i} < 1$) if and only if $E(N_{i,i}) = \frac{1}{1-f_{i,i}} < \infty$.

Using the indicator function, we see that

$$E(N_{i,i}) = \sum_{n=1}^{\infty} E(I\{X_n = i | X_0 = i\}) = \sum_{n=1}^{\infty} P(\{X_n = i | X_0 = i\}) = \sum_{n=1}^{\infty} p_{i,i}^{(n)}$$

Thus, state i is recurrent if and only if

$$\sum_{n=1}^{\infty} p_{i,i}^{(n)} = \infty$$

and otherwise, state i is transient.

Theorem 88. If state i is recurrent, and state i communicates with state j, then state j is recurrent. Further, if state i is transient, and state i communicates with state j, then state j is transient

Proof:

State i is recurrent:

Since i communicates with j, there exists positive integers k and m such that $p_{i,j}^{(k)} > 0$ and $p_{j,i}^{(m)} > 0$. Thus, for any positive integer n, we have

$$p_{j,j}^{(m+n+k)} \geq p_{j,i}^{(m)} p_{i,i}^{(n)} p_{i,j}^{(k)}$$

The above inequality is true since the probability on the left allows for any route between j back to j in $m+n+k$ steps. While the quantity on the right is also the probability of returning to j in $m+n+k$ steps but with the added restrictions of visiting state i at step m, returning to i in n steps, and then traveling to j in k steps. (Note that $p_{i,i}^{(n)}$ could be 0 for some values of n.)

Summing the above inequality over all values of n, we have

$$\sum_{n=1}^{\infty} p_{j,j}^{(n)} \geq \sum_{n=1}^{\infty} p_{j,j}^{(m+n+k)} \geq p_{j,i}^{(m)} p_{i,j}^{(k)} \sum_{n=1}^{\infty} p_{i,i}^{(n)} = \infty$$

Thus, state j is also recurrent.

State i is transient:

This follows easily from the above result. For if j were recurrent, then we know that i would also be recurrent (contradiction to our assumption. ■

Theorem 88 implies the following result:

Theorem 89. If i and j are in the same communication class, then both are either recurrent or both are transient. Further, for an irreducible Markov chain, either all states are recurrent or all states are transient.

For an irreducible Markov chain, if all states are recurrent, then the Markov chain is said to be recurrent; otherwise, the Markov chain is said to be transient.

Theorem 90. An irreducible Markov chain with a finite state space is recurrent.

Proof: If a Markov chain has a finite state space, then not all the states can be transient; otherwise, after a finite number of steps the chain would leave every state never to return and thus have nowhere to go. ■

8.5 Periodicity

The states of a Markov chain can be periodic in terms of return visits. For example, the states in a 1-dimensional random walk are all of period 2 since the return time (in steps) is always an even number.

To define periodicity, we use the function

$$d(x) = \gcd\{n \mid p_{x,x}^{(n)} > 0\}$$

For a given state x, if $d(x) > 1$, the state is said to be **periodic**.

For example, for a state x in a Markov chain $p_{x,x}^{(1)} > 0$ (i.e., there is a positive probability of a return to state x in one step), then state x is aperiodic, since $d(x) = 1$.

Figure 7 depicts a Markov chain all of whose states are of period three. For example, consider state 1. Starting in state 1, one can return in $3,6,9,\ldots$ steps. The GCD of these numbers is 3 and thus state 1 is of period 3.

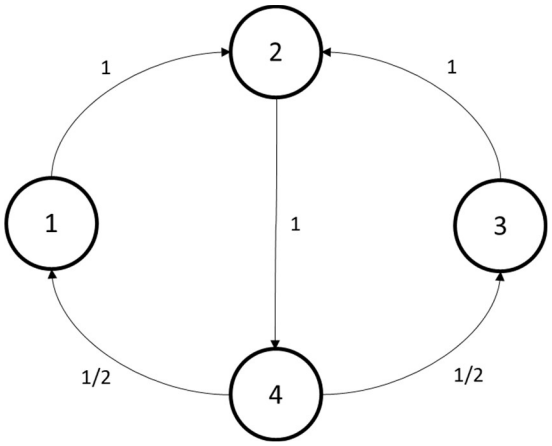


Figure 7. Markov chain with states of period 3

Periodicity, like recurrence and transience, is a class property that is shared by all states in a communication class. In particular, we have the following theorem:

Theorem 91. If states x and y are in the same communication class, then $d(x) = d(y)$.

Proof:

Since x and y communicate, we know there exists positive integers k and m such that

$$p_{x,y}^{(k)} > 0, \qquad p_{y,x}^{(m)} > 0$$

We have that

$$p_{x,x}^{(k+m)} \geq p_{x,y}^{(k)} p_{y,x}^{(m)} > 0$$

since the paths from x to x of length $k + m$ include (are a superset of) the set of paths from x to x going first from x to y in k steps and then from y to x in m steps. Since x has period $d(x)$, $k + m$ must be a multiple of $d(x)$.

Next, let l be any positive integer such that $p_{y,y}^{(l)} > 0$. Using a path argument similar to that in the proof of Theorem 88, we have that

$$p_{x,x}^{(k+l+m)} \geq p_{x,y}^{(k)}\, p_{y,y}^{(l)}\, p_{y,x}^{(m)} > 0$$

Since x has period $d(x)$, $k + l + m$ must be a multiple of $d(x)$.

Thus, $(k + l + m) - (k + m) = l$ must be a multiple of $d(x)$. Since l is any positive integer such that $p_{y,y}^{(l)} > 0$, it must be that $d(x)$ is a divisor of all such l. So, $d(x) \leq d(y)$ by the definition of GCD.

If we reverse the roles of x and y in the above arguments, then we also have $d(y) \leq d(x)$.

Putting the two results together gives us $d(x) = d(y)$. ■

The Markov chain in Figure 8 is irreducible, i.e., every state can be reached by every other state. So, by Theorem 91, we only need to determine the periodicity of one state to determine the periodicity of all the states.

Let's analyze state 4. We have the following paths for returning to state 4:

- $4 \to 3 \to 2 \to 1$ (3 steps)
- $4 \to 3 \to 5 \to 6 \to 7 \to 8 \to 3 \to 2 \to 4$ (8 steps)

There are other return loops, but they are all combinations of the above loops. The GCD of 3 and 8 is 1, and so, $d(4) = 1$. Thus, state 4 is aperiodic and by Theorem 91, all the states of the Markov chain are aperiodic.

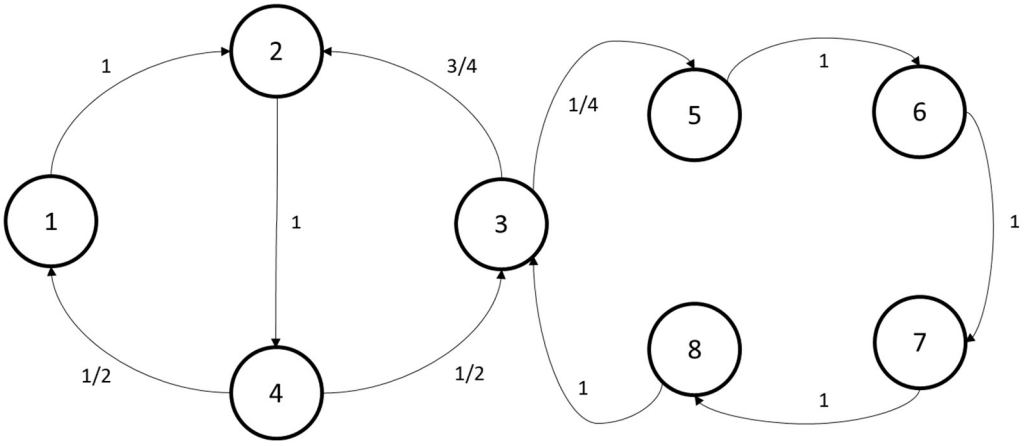


Figure 8. Aperiodic Markov chain

The irreducible Markov chain in Figure 9 is a variation of the Markov chain in Figure 8, with one additional state. With this modification, the Markov chain is now periodic. If we consider state 1, the following are some example loops:

- $1 \to 2 \to 4 \to 1$ (3 steps)
- $1 \to 2 \to 4 \to 3 \to 2 \to 4 \to 1$ (6 steps)
- $1 \to 2 \to 4 \to 3 \to 5 \to 6 \to 7 \to 8 \to 9 \to 3 \to 2 \to 4 \to 1$ (12 steps)

All other loops back to 1 are also multiples of 3 and so, $d(1) = 3$. By Theorem 91, all the states in the Markov chain are of period 3.

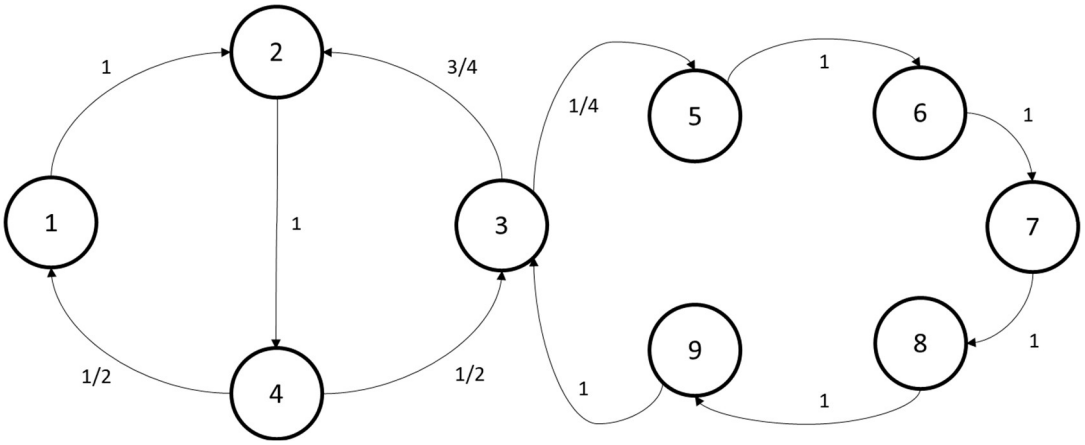


Figure 9. Markov chain with period 3

The transition matrix for the above diagram is as follows:

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

9 Least Squares

He who knows only his own side of the case knows little of that.

John Stuart Mill

9.1 Problem Statement

In this section, we explore the concept of solving over-determined linear systems (i.e., more equations than unknowns) by minimizing the sum of squared errors in the equations. Known as the least squares method, this approach has broad applications across many fields. The technique was independently developed by mathematicians Carl Friedrich Gauss and Adrien-Marie Legendre in the early 19th century.

Assume that we are given a "tall" $m \times n$ matrix A where "tall" means $m > n$. Let x be an $n \times 1$ column vector and b be an $m \times 1$ column vector. In this case, the system of linear equations $Ax = b$ is over-determined. The system of equations $Ax = b$ has a solution only when b is a linear combination of the columns of A. Thus, for some values of b, there is no solution to $Ax = b$. All is not lost, however. We can determine x for such that $e = Ax - b$ (known as the **residual** for $Ax = b$) is of least value. In particular, we determine x to minimize the square of Frobenius norm of $Ax - b$, i.e., find x such that $ee^T = \|Ax - b\|_F^2$ is minimal.

The problem of finding a vector $\hat{x}$ that minimizes $\|Ax - b\|_F^2$ is known as the **least squares problem** and is denoted by

$$\text{minimize } \|Ax - b\|_F^2$$

The input data for the problem are A and b. The quantity to be minimized (i.e., ee^T) is known as the **objective function**. Any vector $\hat{x}$ satisfying $\|A\hat{x} - b\|_F^2 \leq \|Ax - b\|_F^2$ for all other x is a solution to the least squares problem. Such a vector is known as a least squares approximate solution to $Ax = b$. It is important to know that $A\hat{x} \neq b$ unless, as noted earlier, b is a linear combination of the columns of A.

The least squares problem can be cast in terms of the columns of A. If $c_1, c_2, \dots, c_n$ are the columns A then the least squares problem can be viewed as finding a linear combination of the column vectors that is closest to b, i.e.,

$$\text{minimize } \|Ax - b\|_F^2 = \|(x_1c_1 + x_2c_2 + \cdots + x_nc_n) - b\|^2$$

If $\hat{x} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n)^T$ is the solution of the least squares problem, then $A\hat{x} = \hat{x}_1c_1 + \hat{x}_2c_2 + \cdots + \hat{x}_nc_n$ is the closest vector to b among all linear combinations of $c_1, c_2, \dots, c_n$.

A row-based interpretation of the problem is also possible. In this case, let $r_1, r_2, \dots, r_m$ be the rows of matrix A. The components of the residual vector e are

$$e_i = r_ix - b_i, \qquad i = 1,2, \dots, m$$

In this interpretation, the least squares problem is to find x that minimizes the objective function

$$\|Ax - b\|_F^2 = (r_1x - b_1)^2 + (r_2x - b_2)^2 + \cdots + (r_mx - b_m)^2$$

9.2 Solution

Our problem is to minimize the objective function $f(x) = ee^T = (Ax - b)(Ax - b)^T$ under the assumption that A is of full column rank, i.e., the columns of A are linearly independent.

Using the product rule to differentiate $f(x)$, we get

$$df(x) = e\,de^T + de\,e^T = e\,de^T + (de\,e^T)^T = 2e\,de^T = 2e^T\,de = 2e^T(A\,dx)$$

Hence, the minimum is achieved when $e^T A = 0$, i.e., when

$$(x^T A^T - b^T)A = 0$$

$$x^T A^T A = b^T A$$

$$A^T Ax = A^T b$$

$$\boldsymbol{x = \left(A^T A\right)^{-1} A^T b = A^+ b}$$

Recall that $A^+ = (A^T A)^{-1} A^T$ is the Moore-Penrose pseudo inverse of A (see Section 3.12) which exists since we assumed A was of full column rank.

This is the most straightforward approach to solving the least square minimization problem, i.e., compute $A^T b$, compute $(A^T A)^{-1}$ and then multiple the two. However, for larger size matrices, this is not the most efficient solution.

A more efficient solution involves the QR decomposition of matrix A. Under our assumption that the columns of A are linearly independent, we know that the QR decomposition of A exists (see Section 3.18). So, we write $A = QR$. Next, we write A^+ in terms of R and Q as follows:

$$A^+ = (A^T A)^{-1} A^T$$

$$= (R^T Q^T QR)^{-1} (QR)^T$$

$$= R^{-1} Q^{-1} (Q^T)^{-1} (R^T)^{-1} R^T Q^T$$

$$= R^{-1} Q^{-1} = R^{-1} Q^T$$

In the above, we made use of the fact that Q is an orthogonal matrix, i.e., $Q^{-1} = Q^T$.

So, the minimum least squares solution can be written in the form $x = R^{-1} Q^T b$.

In terms of an algorithm, the following steps are to be used:

1. Compute the QR factorization for A.
2. Compute $Q^T b$.
3. Solve the equation $Rx = Q^T b$. This step is known as “back substitution”.

9.3 Comparison of Solution Approaches

In what follows, we compare the pseudo-inverse approach to the approach based on the QR decomposition of A. The "Big O notation" [49] is used in the explanations.

- Existence and Uniqueness:
  - Both methods provide the same unique solution when A is full column rank.
  - When A is rank-deficient, the pseudo-inverse gives the minimum-norm least squares solution, while QR can be adapted (e.g., via column pivoting) to handle rank deficiency.
- Numerical Stability:
  - QR decomposition is numerically stable because it avoids forming A^TA, which can square the condition number of A. (The condition number of a matrix measures how sensitive the solution of a linear system $\mathrm{Ax} = \mathrm{b}$ is to small changes in A or b.)
  - The pseudo-inverse approach via $(A^TA)^{-1}A^T$ can be unstable for ill-conditioned A due to the inversion of A^TA.
- Computational Cost:
  - Pseudo-inverse: Computing A^TA is $O(mn^2)$, inverting A^TA is $O(n^3)$, and multiplying by A^Tb is $O(n^2)$.
  - QR decomposition: Computing $A = QR$ is $O(mn^2)$ (similar to forming A^TA), but solving $Rx = Q^Tb$ is $O(n^2)$.
  - For large m, both methods are dominated by $O(mn^2)$, but QR is more accurate.
- Implementation:
  - Pseudo-inverse is straightforward but can be numerically hazardous.
  - QR decomposition is more involved but is the method of choice in many numerical libraries.

9.4 Example

Let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

We should check that the columns of A are linearly independent, which we can do with a few elementary column operations, i.e.,

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

We will determine the value of x such that $\|Ax - b\|_F^2$ is at a minimum, using the pseudo inverse method.

First, we need to compute A^TA and A^Tb

$$A^TA = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

$$A^Tb = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \end{bmatrix}$$

Next, we need to solve $A^TAx = A^Tb$ which can be written as

$$3x_1 + 6x_2 = 5$$

$$6x_1 + 14x_2 = 11$$

Solving the above equations, we get $x_1 = \frac{2}{3}$ and $x_2 = \frac{1}{2}$.

We have that $Ax - b = \begin{bmatrix} \frac{7}{6} \\ \frac{5}{3} \\ \frac{13}{6} \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} \\ -\frac{1}{3} \\ \frac{1}{6} \end{bmatrix}$ and $\|Ax - b\|_F^2 = \frac{1}{36} + \frac{1}{9} + \frac{1}{36} = \frac{6}{36} = \frac{1}{6}$

9.5 Generalization

We can extend the least squares problem if we take x to a matrix rather than a vector, i.e., let X be an $n \times k$ matrix. The new problem is to minimize $\|AX - B\|_F^2$, where A is an $m \times n$ matrix and B is an $m \times k$ matrix. If we express $AX - B$ in terms of the corresponding column vectors, then the problem reduces to

$$\|AX - B\|_F^2 = \|Ax_1 - b_1\|_F^2 + \|Ax_2 - b_2\|_F^2 + \cdots + \|Ax_k - b_k\|_F^2$$

where x_j is column j of X and b_j is column j of B.

In words, the objective function of the more general problem is just the sum of k instances of the initial problem where x and b were vectors. We can solve each of the component problems separately, i.e., minimize $\|Ax_j - b_j\|_F^2,\ j = 1,2, \dots, k$. Based on our previous work, the solution to each component problem is $x_j = A^*b_j, j = 1,2, \dots, k$. The solution to the overall problem is

$$X = [\, x_1 \quad x_2 \quad \dots \quad x_k \,]$$

$$= [\, A^*b_1 \quad A^*b_2 \quad \dots \quad A^*b_k \,]$$

$$= A^*[\, b_1 \quad b_2 \quad \dots \quad b_k \,]$$

$$= A^*B$$

9.6 Data Fitting

The least squares model is a fundamental framework used in simple data fitting problems where you have n input variables and one output variable.

In terms of problem setup, there are m data points (observations), each with n input features. These can be represented as rows in matrix A, where each row corresponds to one observation, and each column corresponds to a feature (including an intercept/bias term if needed).

$$A = \begin{bmatrix} 1 & a_{11} & a_{12} & \dots & a_{1n} \\ 1 & a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

The first column accounts for the intercept/bias term in linear regression.

The corresponding outputs (for each observation) are collected into vector b.

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$$

The model parameters (weights) are represented by the vector $x = (x_0, x_1, x_2, \dots, x_n)^T$. The key point here is that a linear relationship is assumed among the inputs and outputs, i.e., $b \cong Ax$ and that x is the minimum solution to $\|Ax - b\|_F^2$ which we already know how to compute from our previous discussion.

Once we compute x from the initial data points, we can use x to model a function $f: R^n \to R$ that takes an input vector $d = (d_1, d_2, \dots, d_n)$ and provides an approximation for the output, i.e.,

$$f(d_1, d_2, \dots, d_n) = x_0 + d_1 x_1 + d_2 x_2 + \cdots + d_n x_n$$

Let's take a simple example where we have but 1 feature and 4 observations, shown in the table below.

Observation (i)	**Input (a_{i1})**	**Output (b_i)**
1	1	2
2	2	3
3	3	5
4	4	6

Converting this to the linear least squares model, we have

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \qquad b = \begin{bmatrix} 2 \\ 3 \\ 5 \\ 6 \end{bmatrix}$$

Doing the computations for the pseudo inverse solution to the least square problem, we have

$$A^T A = \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}, \qquad A^T b = \begin{bmatrix} 16 \\ 47 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 16 \\ 47 \end{bmatrix}$$

Solving the above equation for x, get our answer, i.e.,

$$x = \left(\frac{1}{2}, \frac{7}{5}\right)^T$$

For another data point d, we can approximate its value using $f(d) = x_0 + x_1 d = \frac{1}{2} + \frac{7}{5}d$. For example, $f(4) = \frac{1}{2} + \frac{28}{5} = 6.1$ (compare to the initial data set where $4 \to 6$). Of course, the true value is in prediction, e.g., $f(5) = \frac{1}{2} + 7 = 7.5$.

Further, we can compute the residual and norm of the residual.

$$Ax - b = \begin{bmatrix} 1.9 \\ 3.3 \\ 4.7 \\ 6.1 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} -.1 \\ .3 \\ -.3 \\ .1 \end{bmatrix}, \qquad \|Ax - b\|_F^2 = .2$$

Acronyms and Symbols

If two wrongs don't make a right, try three. – Laurence J. Peter

∀ - for every

∋ - such that

∃ - there exists

∄ - there does not exist

⊗ - Kronecker product

$a \in B$ –> a is an element of set B

$a \notin B$ –> a is not an element of set B

$A \subset B$ –> A is a proper subset of B, i.e., A cannot equal B

$A \subseteq B$ –> A is a subset of B and could possibly equal B

$X \backslash Y$ –> the elements of set X with any elements in common with set Y removed

$\mathbb{C}$ - the set of complex numbers

δ_j^i – Kronecker delta symbol

$\epsilon_{a_1 a_2 \ldots a_n}$ – Levi-Civita symbol

$\mathcal{E}_n$ – standard basis for the vector space $\mathbb{R}^n$

$\mathbb{Q}$ - the set of rational numbers, i.e., fractions

$\mathbb{N}$ - the set of natural numbers, i.e., 1,2,3, …

$\mathbb{R}$ - the set of real numbers

$\mathbb{Z}$ - the set of integers, i.e., …-2,-1,0,1,2,...

LU – Lower-Upper (decomposition)

ML – Machine Learning

NMF – Non-negative Matrix Factorization

RGB – Red Green Blue

SVD – Singular Value Decomposition

References

The whole problem with the world is that fools and fanatics are always so certain of themselves, and wiser people so full of doubts.

Bertrand Russell

[1] *Gaussian elimination*, Wikipedia, https://en.wikipedia.org/wiki/Gaussian_elimination, accessed on 31 March 2025.

[2] Fratini, S., *Mathematical Thinking (2nd edition)*, self-published on Amazon, Paperback: https://www.amazon.com/dp/B0CL34FRP1, Electronic (free): https://github.com/sfratini33/art-of-managing-things-external/blob/master/free_books/MathThink-II.pdf.

[3] *Row echelon form*, Wikipedia, https://en.wikipedia.org/wiki/Row_echelon_form, accessed on 31 March 2025.

[4] *Elementary Matrix*, Wikipedia, https://en.wikipedia.org/wiki/Elementary_matrix, accessed on 1 April 2025.

[5] Angina Seng (https://math.stackexchange.com/users/436618/angina-seng), *Why do row operations not change the column rank?*, URL (version: 2017-11-10): https://math.stackexchange.com/q/2513360.

[6] *Invertible matrix*, Wikipedia, https://en.wikipedia.org/wiki/Invertible_matrix, accessed on 2 April 2025.

[7] Margalit, D., Rabinoff, J., *Interactive Linear Algebra*, Section 3.6, The Invertible Matrix Theorem, https://textbooks.math.gatech.edu/ila/invertible-matrix-thm.html, accessed on 2 April 2025.

[8] Beezer, R.A., *A First Course in Linear Algebra*, Theorem NME4 Nonsingular Matrix Equivalences, http://linear.pugetsound.edu/html/section-CRS.html, accessed on 2 April 2025.

[9] *Composite of Bijections is Bijection*, ProofWiki, https://proofwiki.org/wiki/Composite_of_Bijections_is_Bijection, accessed on 3 April 2025.

[10] *Linear Algebra/Any Matrix Represents a Linear Map*, Wikibooks, https://en.wikibooks.org/wiki/Linear_Algebra/Any_Matrix_Represents_a_Linear_Map, accessed on 11 April 2025.

[11] *Linear Algebra/Changing Representations of Vectors*, Wikibooks, https://en.wikibooks.org/wiki/Linear_Algebra%2FChanging_Representations_of_Vectors , accessed on 14 April 2025.

[12] *Euler's formula*, Wikipedia, https://en.wikipedia.org/wiki/Euler%27s_formula, accessed on 14 April 2025.

[13] *Matrix equivalence*, Wikipedia, https://en.wikipedia.org/wiki/Matrix_equivalence, accessed on 15 April 2025.

[14] Schilling, A., Nachtergaele, B. and Lankham, I., *Linear Algebra*, LibreTexts™, The Gram-Schmidt Orthogonalization procedure, https://math.libretexts.org/Bookshelves/Linear_Algebra/Book%3A_Linear_Algebra_(Schilling_Nachtergaele_and_Lankham)/09%3A_Inner_product_spaces/9.05%3A_The_Gram-Schmidt_Orthogonalization_procedure, accessed on 18 April 2025.

[15] *Linear Algebra/Exploration*, Wikibooks, https://en.wikibooks.org/wiki/Linear_Algebra/Exploration, accessed on 18 April 2025.

[16] *Leibniz formula*, Wikipedia, https://en.wikipedia.org/wiki/Leibniz_formula_for_determinants, accessed on 19 April 2025.

[17] *Linear Algebra/Determinants Exist*, Wikibooks, https://en.wikibooks.org/wiki/Linear_Algebra/Determinants_Exist, accessed on 19 April 2025.

[18] Kuttler, K., *A First Course in Linear Algebra*, LibreTexts™, https://math.libretexts.org/Bookshelves/Linear_Algebra/A_First_Course_in_Linear_Algebra_(Kuttler), accessed on 19 April 2025.

[19] Wikipedia, *Properties of the determinant in relation to other notions*, https://en.wikipedia.org/wiki/Determinant#Properties_of_the_determinant_in_relation_to_other_notions, accessed on 10 August 2025.

[20] *Linear Algebra/The Permutation Expansion*, Wikibooks, https://en.wikibooks.org/wiki/Linear_Algebra/The_Permutation_Expansion, accessed on 21 April 2025.

[21] *Nilpotent matrix*, Wikipedia, https://en.wikipedia.org/wiki/Nilpotent_matrix, accessed on 24 April 2025.

[22] Hefferon, J., *Linear Algebra (4th Edition)*, https://leanpub.com/linalgebra.

[23] *Definite matrix*, Wikipedia, https://en.wikipedia.org/wiki/Definite_matrix, accessed on 21 May 2025.

[24] *Minor*, Wikipedia, https://en.wikipedia.org/wiki/Minor_(linear_algebra), accessed on 21 May 20215.

[25] Huber, K., *Math 26: Introduction to Linear Algebra*, Section 2.6, LibreTexts™, https://math.libretexts.org/Courses/Irvine_Valley_College/Math_26%253A_Introduction_to_Linear_Algebra, accessed on 23 May 2025.

[26] Multilinear map, Wikipedia, https://en.wikipedia.org/wiki/Multilinear_map, accessed on 15 June 2025.

[27] *Kronecker delta*, Wikipedia, https://en.wikipedia.org/wiki/Kronecker_delta, accessed on 11 June 2025.

[28] *Adjugate matrix*, Wikipedia, https://en.wikipedia.org/wiki/Adjugate_matrix#Properties, accessed on 23 May 2025.

[29] *Permutation matrix*, Wikipedia, https://en.wikipedia.org/wiki/Permutation_matrix, accessed on 26 May 2025.

[30] Quixotic (https://math.stackexchange.com/users/2109/quixotic), Proving: "The trace of an idempotent matrix equals the rank of the matrix", URL (version: 2022-03-09): https://math.stackexchange.com/q/101512

[31] *Block matrix*, Wikipedia, https://en.wikipedia.org/wiki/Block_matrix, accessed on 28 May 2025.

[32] *LU decomposition*, Wikipedia, https://en.wikipedia.org/wiki/LU_decomposition, accessed on 28 May 2025.

[33] *QR decomposition,* Wikipedia, https://en.wikipedia.org/wiki/QR_decomposition, accessed on 30 May 2025.

[34] Trefethen, L.N. and Bau III, D., *Numerical Linear Algebra*, Society for Industrial and Applied Mathematics (SIAM), 1997.

[35] *Singular value decomposition,* Wikipedia, https://en.wikipedia.org/wiki/Singular_value_decomposition, accessed 30 May 2025.

[36] *Inverse function theorem*, Wikipedia, https://en.wikipedia.org/wiki/Inverse_function_theorem, accessed 31 May 2025.

[37] Oswaldo Rio Branco de Oliveira, *The Implicit and the Inverse Function theorems: easy proofs*, arXiv:1212.2066, https://arxiv.org/abs/1212.2066, accessed on 31 May 2025.

[38] *Fréchet derivative*, Wikipedia, https://en.wikipedia.org/wiki/Fr%C3%A9chet_derivative, accessed on 5 June 2025.

[39] *Symmetry of second derivatives*, Wikipedia, https://en.wikipedia.org/wiki/Symmetry_of_second_derivatives#Schwarz's_theorem, accessed on 7 June 2025.

[40] *Levi-Civita symbol*, Wikipedia, https://en.wikipedia.org/wiki/Levi-Civita_symbol, accessed on 22 June 2025.

[41] Broida, J.G., Williamson, S.G., *Comprehensive Introduction to Linear Algebra: Part III – Operators and Tensors*, published under Creative Commons CC0 1.0, 2012. https://cseweb.ucsd.edu/~gill/CILASite/.

[42] *Tensor*, Wikipedia, https://en.wikipedia.org/wiki/Tensor, accessed on 23 June 2025.

[43] Lee, J.M., *Introduction to Smooth Manifolds (2nd Edition)**, Graduate Texts in Mathematics, Springer, 2013.

[44] *Perron–Frobenius theorem*, Wikipedia, https://en.wikipedia.org/wiki/Perron%E2%80%93Frobenius_theorem, accessed on 1 July 2025.

[45] Seneta, E., *Non-negative Matrices and Markov Chains (2nd Edition)*, Springer, 1981.

[46] *Markov property*, Wikipedia, https://en.wikipedia.org/wiki/Markov_property, accessed on 3 July 2025.

[47] *Equivalence relation*, Wikipedia, https://en.wikipedia.org/wiki/Equivalence_relation, accessed on 3 July 2025.

[48] Ross, S.M., *Introduction to Probability Models**, 12th Edition, Academic Press, 2019. An earlier edition can be found at https://ia801508.us.archive.org/35/items/collection-of-mathematics-probability-and-statistics-books/Introduction_to_Probability_Models_Tenth_Edition_Sheldon_M_Ross.pdf

[49] *Big O notation*, Wikipedia, https://en.wikipedia.org/wiki/Big_O_notation, accessed on 7 July 2025.

[50] *Gambler's ruin*, Wikipedia, https://en.wikipedia.org/wiki/Gambler's_ruin, accessed on 6 September 2025.

* Indicates the book or article is available for borrowing from the Internet Archive at https://archive.org/. In some cases, only an earlier edition of a book will be available.

Index of Terms