Review over some Math

Notes on the book 'All the Math You Missed But Need to Know for Graduate School' by Thomas A. Garrity, along with some personal thoughts and excercise solutions.

José Daniel Mejía C.

2026

Contents

Chapter 1

Summary over topics

1.1 Structure of mathematics

1.1.1 Equivalence problems

Several problems or fields of study in math come down to seeing if two or more objects are the 'same', or detecting their differences.

> **Note:-**
> The concept of two objects being the same depends on the field of study.

   Equivalence between two objects can be explained by the allowed Maps or functions between them.

> **Example 1.1.1**
>
> A square and a circle are equivalent in topology; they can be stretched out from one to the other.
> They are not the same in differential geometry; one is smooth (differentiable) and the other is not.

   Invariants are properties that two or more object must satisfy if they are equivalent, however they themselves do not determine equivalence between them.

> **Example 1.1.2**
>
> In topology, two circles are not equivalent to a single circle. The two circles have two connected components, whereas the one only has one connected component. Hence, both objects are not equivalent.
> However, a sphere and a circle have the same number of connected components, but they are not equivalent.

1.1.2 Functions

Functions describe all kinds of situations, both in the real world and in an abstract matter.

   Each field of study or area in mathematics focuses in a specific kind of function. Calculus studies differentiable functions fron $\mathbb{R} \to \mathbb{R}$, for example.

1.2 Topics

1.2.1 Linear Algebra

Linear algebra studies **Linear Transformations** and **Vector Spaces**. A linear transformation has a matrix representation, given a basis for the vector space being considered.

Two or more different matrices can represent the same transformation, but over different choices of bases of the same vector space.

The key theorem in linear algebra explains several equivalent descriptions for when a matrix, and hence the associated transformation, is invertible.

Other important aspect is understanding the occurrence and usefulness of eigenvectors and eigenvalues.

1.2.2 Real Analysis

The definitions of **Limit**, **Continuity**, **Differentiation** and **Integration** are given in terms of ϵ and δ. **Uniform Convergence** of functions is a key part of this topic.

1.2.3 Differentiating Vector-Valued Functions

The **Inverse Function Theorem** explains that a differentiable function $f : \mathbb{R} \to \mathbb{R}^n$ is invertible iff the determinant of its derivative (the Jacobian) is never zero.

It is important to understand what it means for a vector-valued function to be differentiable, and why its derivative must always be a linear map, represented by the Jacobian as a matrix.

Understanding the **Implicit function theorem** and how it relates to the Inverse Function Theorem is also relevant.

1.2.4 Point Set Topology

The concept of **Open Sets** and how a **Topology** can be built in terms of open sets must be understood. Also how a function is said to be **Continuous** in terms of open sets is key.

The standard topology in $\mathbb{R}^n$ and the **Heine-Borel Theorem** is covered, as well as what a metric space is and how a topology is defined using its metric.

1.2.5 Classical Stokes' Theorem

The calculus of vector fields covers the **Curl**, and **Divergence** of vector fields, the **Gradient** of a function and the **Path Integral** along a curve.

The **Divergence Theorem** and **Stokes' Theorem** are presented as classical extensions of the Fundamental Theorem of Calculus, it is relevant to understand why they are indeed extensions of this classic theorem.

1.2.6 Differential Forms and Stokes' Theorem

Manifolds are geometric objects. **Differential k-forms** allows calculus to be done on manifolds. There are several ways to define a manifold. How to think about and define differential k-forms, and how to take their **external derivative** is important.

Differential k-forms and external derivatives can be represented using vector fields, gradients, curls and divergences.

The statement of **Stokes' Theorem** in this context is covered, and special cases over the Divergence Theorem and Stokes' Theorem seen before.

1.2.7   Curvature for Curves and Surfaces

Curvature is a measure of the rate of change of the direction of tangent spaces of several geometric objects.

It is relevant to compute the curvature of curves on a plane, the curvature and torsion of a curve in space, and the two principal curvatures of a surface in space, in terms of the Hessian.

1.2.8   Geometry

Different axioms build several geometries. For example, euclidean geometry assumes that given a line l and a point p not on l, there is exactly only one other line containing p parallel to l. Elliptic and hyperbolic geometries assume different conditions.

Models for hyperbolic geometry, single elliptic geometry and double elliptic geometry are covered.

1.2.9   Countability and the Axiom of Choice

The concept of a **Countably Infinite** set is presented. For example, the Integers and Rationals are, while the Real numbers are not.

The statement of **The Axiom of Choice** and its (sometimes weird) equivalences should be covered.

1.2.10   Elementary Number Theory

An introduction to **Modular Arithmetic** is given. Importan explanations as why there are infinitely many primes, what is a **Diophintine Equation**, what is the **Euclidean Algorithm** and how it is linked to **Continuous Fractions** are covered.

1.2.11   Algebra

Groups and their properties should be studied, as well as understanding how these represent geometric symmetries. Basics over Groups, **Rings** and **Fields** are covered.

Galois Theory explains how finding the roots of polynomials and finite group theory are linked. **Representation Theory** is also covered, establishing a relation between abstract groups and groups of matrices.

1.2.12   Algebraic Number Theory

Definitions and examples of **Algebraic Number Fields** are presented. Each of these have themselves an analog of the integers, but not always do these integers have a unique factorization.

1.2.13   Complex Analysis

Complex Analysis presents what it means for a function $f : U \subseteq \mathbb{C} \to \mathbb{C}$ to be **Analytic**, and equivalent conditions that imply analyticity; featuring the **Cauchy-Riemann Equations** and the **Cauchy Integral Formula**.

An analytic function $f : U \to \mathbb{C}$ such that $f'(z_0) \neq 0$, then at z_0 the function f is **Conformal**, (angle preserving), when viewed as a function from $\mathbb{R}^2$ to $\mathbb{R}^2$.

1.2.14   Analytic Number Theory

The definition of the **Riemann Zeta Function**, its relation to prime numbers, and the statement of the **Riemann Hypothesis** are presented.

1.2.15   Lebesgue Integration

The concepts of **Lebesgue Measure** and **Lebesgue Integration** are presented, as well as the **Lebesgue Dominating Convergence Theorem**. It is important to cover what a set of measure zero is, and what the

'almost everywhere' notion means.

1.2.16 Fourier Analysis

It is important to understand how to find the **Fourier Series** of a periodic function, the **Fourier Integral** of a function, the **Fourier Transform**, and how they all relate to **Hilbert Spaces**. Their use in helping solve differential equations is also relevant.

1.2.17 Differential Equations

Several real world problems can be modeled by systems of differential equations. Almost all **Ordinary Differential Equations** have solutions (and simple formulas for them). In contrast, **Partial Differential Equations** do not have general solutions, only in very specific cases; furthermore, sometimes it is not know if there even exists a solution for them.

Important classes of partial differential equations are the heat equation, the wave equation and the Laplacian.

1.2.18 Combinatorics and Probability Theory

Elementary combinatorics and probability theory are covered. The **Binomial Theory** for polynomials is very useful for computations, and occasionally appears as a steps for proofs in several branches of mathematics.

Sample Spaces and **Random Variables**, **Expected Value** and **Variance** are key concepts. The **Central Limit Theorem** links integral calculus to the subject.

1.2.19 Algorithms

Understanding what the **Complexity** of an algorithm is, and the question **P=NP** is relevant. Some graph theory and algorithms over graphs are covered, as well as why graphs are relevant in the representation of problems.

Numerical Analysis algorithms and methods are important for computing approximations for problems in mathematics.

1.2.20 Category Theory

Category Theory is a method for thinking about mathematics. Each area of mathematics could be put in the language of categories, being described by its relevant mathematical objects and the morphisms between them, with the results being described by diagrams of arrows.

Chapter 2

Linear Algebra

> **Note:-**
> The basic mathematical object is a **Vector Space**, the basic map is a **Linear Transformation**, and the objective is to find equivalent conditions for **Invertibility of Matrices**

Linear Algebra is one of the most relevant (if not the most relevant from my experience) area of mathematics. It allows us to solve systems of linear equations. However, as interesting as solving linear equations might be, the power of linear algebra comes when the abstraction of several different mathematical objects can be represented as vector spaces, hence they can be modified and mapped to and from by linear transformations.

Here, solving the systems (and knowing when a solution for them exists) becomes way more interesting.

2.1 Basic concepts

$\mathbb{R}^n$ is the most basic example of a vector space. It is the set of all n-tuples of real numbers

$$\{(x_1, \ldots, x_n) \ : \ x_i \in \mathbb{R}\}$$

Each n-tuple is usually called a **vector**, and two can be added to get another one

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

or they could be multiplied by a number λ, usually called a **scalar**.

$$\lambda (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$$

The natural map from some $\mathbb{R}^n$ to a $\mathbb{R}^m$ is matrix multiplication by a $n \times m$ matrix. Matrix multiplication is **linear**.

Example 2.1.1 (Left Matrix Multiplication)

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \\ 2 \end{pmatrix}$$

Where $\begin{pmatrix} 2 \\ 3 \end{pmatrix} \in \mathbb{R}^2$ and $\begin{pmatrix} 5 \\ 5 \\ 2 \end{pmatrix} \in \mathbb{R}^3$.

Systems of linear equations can be expressed as a left matrix multiplication. Our goal is to find the correct value for each x_i.

$$a_{11}x_1 + \cdots + a_{1n}x_n = b_1$$
$$\vdots$$
$$a_{m1}x_1 + \cdots + a_{mn}x_n = b_m$$

This system becomes $Ax = b$, where

$$A = \begin{pmatrix} a_{11} \ldots a_{1n} \\ \vdots \\ a_{m1} \ldots a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

When $m > n$ (more equations than unknown variables) there are usually no solutions to the system (except in some linearly dependent cases). Conversely, when $m < n$, we expect to have several solutions.

The most common situation though, is when $m = n$. In this case, finding if an inverse to A, A^{-1} exists is our objective, because then finding the vector x is very straightforward.

$$Ax = b \quad \implies \quad A^{-1}Ax = A^{-1}b \quad \implies \quad x = A^{-1}b$$

Hence, finding when a square matrix has an inverse is the same as determining when an $n \times n$ system of linear equations has a solution.

2.2 Vector Spaces and Linear Transformations

Definition 2.2.1 Vector Space

A set V is a **Vector Space** over a field K (usually $\mathbb{R}$ or $\mathbb{C}$) if there are maps:

1. $\cdot : K \times V \to V$ multiplication by scalar, av for $a \in K, \ v \in V$

2. $+ : V \times V \to V$ vector sum, $v + u$ for $v, u \in V$

with the properties:

(a) $\exists 0 \in V \ : \ 0 + v = v \ \forall v \in V$

(b) $\forall v \in V \ \exists (-v) \in V \ : (-v) + v = 0$

(c) $\forall v, u \in V \ v + u = u + v$

(d) $\forall v, u \in V, \ \forall a \in K \ a(u + v) = au + av$

(e) $\forall v \in V, \ \forall a, b \in K \ a(bv) = (ab)v$

(f) $\forall v \in V, \ \forall a, b \in K \ (a + b) = av + bv$

(g) $\exists 1 \in K \ \forall v \in V \ 1v = v$

These are all usual properties easily recognizable in the vector space $\mathbb{R}^n$ over the field $\mathbb{R}$.

Definition 2.2.2 Linear Transformation

A **Linear Transformation** T is a map from any vector space V to any other vector space W, such that

$$T(a_1 u + a_2 v) = a_1 T(u) + a_2 T(v)$$

7

> for any two vectors and scalars.

Left matrix multiplications as seen before is a linear transformation.

Definition 2.2.3 Subspace

A subset $U \subset V$ is a **Subspace** if it itself is a vector space.

It is way easier to check this condition however,

Proposition 2.2.1 $U \subset V$ is a subspace if it contains the zero vector and it's closed under addition and scalar multiplication

Given a linear transformation T from a vector space V to another W, two natural subspaces are the kernel and image.

Definition 2.2.4 Kernel & Image of a Linear Map

Given a linear map $T : V \to W$, its **Kernel** is

$$\ker(T) = \{v \in V \mid T(v) = 0 \in W\}$$

and its **Image** is

$$\text{Im}(T) = \{T(v) \in W \mid v \in V\}$$

Clearly both are subsets of V and W respectively, and closed under addition and multiplication by scalars.

Example 2.2.1 (Derivative as linear map)

In the vector space $C^k[0,1]$, the set of all real-valued functions with domain over $[0,1]$ such that the kth derivative exists and is continuous, the derivative is a linear transformation from $C^k[0,1]$ to $C^{k-1}[0,1]$.

$$\frac{d}{dx} : C^k[0,1] \to C^{k-1}[0,1]$$

Then the kernel of this linear transformation are all constant functions.
For example, in the differential equation

$$f'' + 3f' + 2f = 0$$

we can define the linear map

$$T = \frac{d^2}{dx^2} + 3\frac{d}{dx} + 2I \ : \ C^2[0,1] \to C^0[0,1]$$

Hence, the solution lies in $\ker(T)$.

2.3 Bases, Dimension, and Linear Transformations as Matrices

Definition 2.3.1 Basis

A set of vectors $\{v_1, \ldots, v_n\}$ form a **Basis** for a vector space V if for any $v \in V$ there are unique scalars $a_1, \ldots, a_n \in K$ such that

$$a_1 v_1 + \cdots + a_n v_n = v$$

Definition 2.3.2 Dimension

The **Dimension** of a vector space V, $\dim(V)$ is the number of elements in a basis.

Dimension for a vector space is always a fixed number, hence, it is necessary to prove that every basis for a vector space has the same number of elements.

Theorem 2.3.1

All bases of a vector space have the same number of elements

To prove this theorem, we will need the following definitions:

Definition 2.3.3 Linear Independence

A set of vectors $\{v_1, \ldots, v_n\}$ is **Linearly Independent** if whenever

$$a_1 v_1 + \cdots + a_n v_n = 0$$

implies

$$a_1 = \cdots = a_n = 0$$

Definition 2.3.4 Span

A set of vectors $\{v_1, \ldots, v_n\}$ span the vector space V if any $v \in V$ can be expressed as a linear combination of vectors in $\{v_1, \ldots, v_n\}$.

Then, a basis is a linearly independent set of vectors that spans the vector space associated to it.

This restricts any basis for a vector to have the same elements as any other.

2.3.1 Linear Transformations as Matrices

To see how a linear transformation between two **finite-dimensional** vector spaces can be represented as a matrix, we first need to fix a basis for both. This is not trivial, as when working with more difficult spaces a basis is not always given.

Let us have V and W, vector spaces over K. First, given a linear transformation $T : V \to W$, we need to pick a basis for each vector space; $\beta = \{v_1, \ldots, v_n\}$ for V, and $\gamma = \{w_1, \ldots, w_m\}$ for W. Then, any vector $v \in V$ can be expressed as

$$v = a_1 v_1 + \cdots + a_n v_n \qquad \text{for unique } a_1, \ldots, a_n \in K$$

then, a one to one correspondence between V and K^n is made, because for every $v \in V$ we can pick a unique column vector in K^n, and every column vector in K^n has a corresponding vector in V.

$$v = a_1 v_1 + \cdots + a_n v_n \in V \quad \text{maps to} \quad [v]_\beta = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in K^n$$

> **Note:-**
>
> This is, in fact, an **isomorphism** $[.]_\beta : V \to K^n$, with $n = \dim(V)$. However, this isomorphism is **not** canonical, which means it depends on the basis given to the vector space. Which in short means if we don't have a basis there is no evident isomorphism.

the same can be done for any $w \in W$:

$$w = b_1 w_1 + \cdots + b_m w_m \in W \quad \text{maps to} \quad [w]_\gamma = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m$$

Then, each vector in any vector space will be referred to by its coordinates in the Cartesian product of its field, meaning its corresponding unique column vector under the given basis.

Now, each v_i in the basis for V has a corresponding image in W. $T(v_i) = w$, which in itself can be represented by a linear combination of elements in $\{w_1, \ldots, w_m\}$. Then,

$$T(v_i) = a_{1i} w_1 + \cdots + a_{mi} w_m, \quad \text{where } a_{1i}, \ldots, a_{mi} \in K$$

Given any vector $v \in V$ where $v = a_1 v_1 + \cdots + a_n v_n$, then

$$
\begin{aligned}
T(v) &= T(a_1 v_1 + \cdots + a_n v_n) \\
&= T(a_1 v_1) + \cdots + T(a_n v_n) \\
&= a_1 T(v_1) + \cdots + a_n T(v_n) \\
&= a_1 (a_{11} w_1 + \cdots + a_{m1} w_m) + \cdots + a_n (a_{1n} w_1 + \cdots + a_{mn} w_m) \\
&= a_1 a_{11} w_1 + \cdots + a_1 a_{m1} w_m + \cdots + a_n a_{1n} w_1 + \cdots + a_n a_{mn} w_m \\
&= (a_1 a_{11} + \cdots + a_n a_{1n}) w_1 + \cdots + (a_1 a_{m1} + \cdots + a_n a_{mn}) w_m
\end{aligned}
$$

Following the correspondence between W and K^m, this vector corresponds to

$$[T(v)]_\gamma = \begin{pmatrix} a_1 a_{11} + \cdots + a_n a_{1n} \\ \vdots \\ a_1 a_{m1} + \cdots + a_n a_{mn} \end{pmatrix}$$

which is, in fact, the matrix multiplication $[T]_\beta^\gamma [v]_\beta$, where

$$[T]_\beta^\gamma = \begin{pmatrix} | & & | \\ [T(v_1)]_\gamma & \cdots & [T(v_n)]_\gamma \\ | & & | \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in M_{m \times n}(K) \quad \text{and} \quad [v]_\beta = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

So, if $b_j = a_1 a_{j1} + \cdots + a_n a_{jn}$, then

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

And $[T]_\beta^\gamma$ is the matrix representation of T.

If $T : V \to V$ is a linear map from one vector space to itself, then its matrix representation will be an $n \times n$ square matrix.

> **Note:-**
>
> A linear transformation's matrix representation **will** change if the basis chosen for any of the spaces changes. A key result is identifying if two matrix representations refer to the same linear map, just with different basis being chosen.

Example 2.3.1 (Linear Map in $\mathbb{R}^n$)

Let us consider the linear map $f : \mathbb{R} \to \mathbb{R}^3$, defined by:

$$f(x) = \begin{pmatrix} x \\ 3x \\ \frac{1}{2}x \end{pmatrix}$$

And let us choose the basis

$$\beta = \{6\} \text{ for } \mathbb{R}, \quad \text{and} \quad \gamma = \left\{ \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \text{ for } \mathbb{R}^3$$

Calculating the transform on the element of β,

$$f(6) = \begin{pmatrix} 6 \\ 3 \times 6 \\ \frac{1}{2} \times 6 \end{pmatrix} = \begin{pmatrix} 6 \\ 18 \\ 3 \end{pmatrix}$$

And we express it as a linear combination of elements in γ,

$$f(6) = \begin{pmatrix} 6 \\ 18 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} + \frac{15}{2} \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{which maps to} \quad [f(6)]_\gamma = \begin{pmatrix} 1 \\ \frac{15}{2} \\ 3 \end{pmatrix}$$

Meaning the linear map f has a matrix representation:

$$[f]_\beta^\gamma = \begin{pmatrix} 1 \\ \frac{15}{2} \\ 3 \end{pmatrix}$$

Confirming our calculations, using 2 as an example:

$$f(2) = \begin{pmatrix} 2 \\ 3 \times 2 \\ \frac{1}{2} \times 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 1 \end{pmatrix}$$

And using matrix representation,

$$2 = \frac{1}{3}6 \implies [2]_\beta = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad \text{then,} \quad [f]_\beta^\gamma [2]_\beta = \begin{pmatrix} 1 \\ \frac{15}{2} \\ 3 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{5}{2} \\ 1 \end{pmatrix}$$

which when recovering from coordinates,

$$f(2) = \frac{1}{3} \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} + \frac{5}{2} \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 1 \end{pmatrix}$$

Note:-

The last step, 'recovering' our vector, was in fact just a change of basis from the arbitrary one we chose to the usual canonical one. This is just another linear map.

It is the **identity map** from our chosen basis γ to the canonical (standard) basis E.

$$[v]_E = [I]_\gamma^E [v]_\gamma$$

This way of representing the identity map can be used to change any vector representation from one basis to any other basis.

2.4 The Determinant

We will work over square matrices over the **real** numbers.

> **Definition 2.4.1** Determinant of a 1×1 Matrix
>
> The **Determinant** of a 1×1 matrix (a) is the function
> $$\det(a) = a$$

> **Definition 2.4.2** Determinant of a Matrix
>
> The **Determinant** of an $n \times n$ matrix A is the function
> $$\det(A) = \sum_{k=1}^{n}(-1)^{k+1}a_{1k}\det(A_{1k})$$
> where A_{ij} is the resulting matrix when deleting the ith row and the jth column of A.

This yields the familiar way to define the 2×2 determinant,

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}\det(a_{22}) - a_{12}\det(a_{21}) = a_{11}a_{22} - a_{12}a_{21}$$

This is a valid way to define this crucial function in linear algebra, but really fails to convey how it works; or most importantly why it is so relevant.

To define it the second way, we will need the notation

$$A = (A_1, \ldots, A_n)$$

where A_i denotes the ith column of A.

> **Definition 2.4.3** Determinant of a Matrix
>
> The **Determinant** of A is defined as the **unique** real-valued $\det : M_{n \times n}(\mathbb{R})$ function that satisfies, given a scalar λ:
>
> (a) $\det(A_1, \ldots, \lambda A_k, \ldots, A_n) = \lambda \det(A_1, \ldots, A_n)$
>
> (b) $\det(A_1, \ldots, A_i + \lambda A_i, \ldots, A_n) = \det(A_1, \ldots, A_n)$ for $k \neq i$
>
> (c) $\det(I) = 1$

In short, it is the only function where you can

(a) 'Factor' out a number from any column onto the determinant.

(b) Add or subtract any column to another without changing the result.
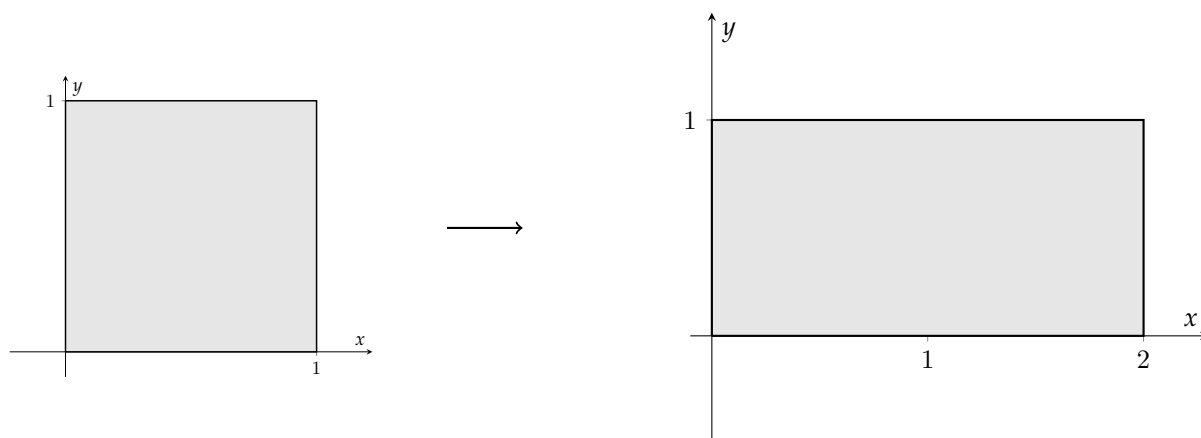
(c) The identity matrix has determinant 1.

The third way of defining it is very geometric, but not too rigorous. However I feel it is very intuitive.

> **Definition 2.4.4** Determinant of a Matrix
>
> The **Determinant** of a matrix is the signed volume of the image of the unit cube.

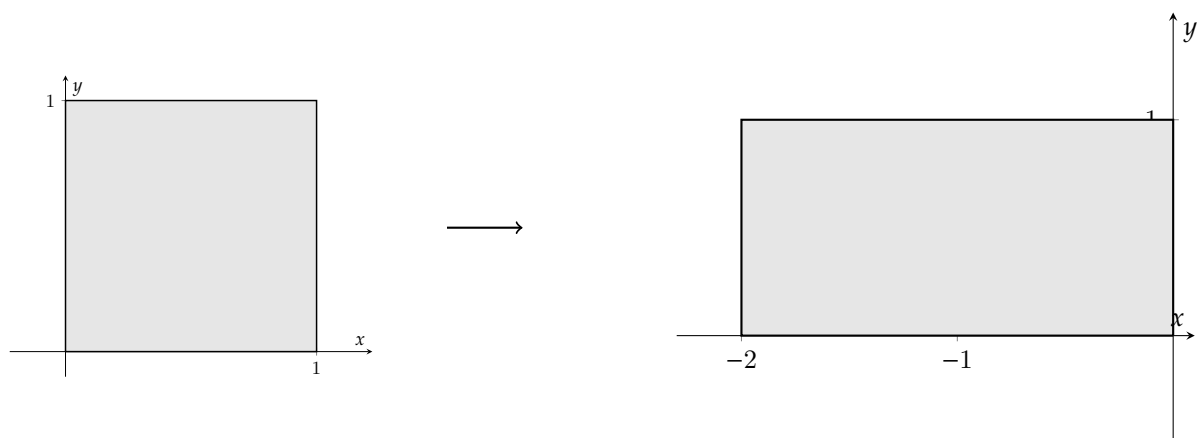To understand let us see an example

The linear map $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ takes the unit square to

Since the area has doubled,

$$\det(A) = 2$$

Now, consider the matrix $A = \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}$

Since the volume doubled, but the orientation of the edges changed, we have

$$\det(A) = -2$$

One additional property of the determinant is

Lemma 2.4.1 Determinant as a Multiplicative Function

If $A, B \in M_{n \times n}(\mathbb{R})$, then

$$\det(AB) = \det(A)\det(B)$$

2.5   The Key Theorem of Linear Algebra

> **Theorem 2.5.1** Key Theorem
>
> Let $A \in M_{n \times n}$, then the following are equivalent:
>
> 1. A is **Invertible**
>
> 2. $\det(A) \neq 0$
>
> 3. $\ker(A) = 0$
>
> 4. If $\boldsymbol{b}$ is a column vector in $\mathbb{R}^n$, then there is a unique column vector $\boldsymbol{x}$ in $\mathbb{R}^n$ satisfying $A\boldsymbol{x} = \boldsymbol{b}$
>
> 5. The columns of A are linearly independent $n \times 1$ column vectors
>
> 6. The rows of A are linearly independent $n \times 1$ column vectors
>
> 7. A^t is invertible
>
> 8. All the eigenvalues of A are non-zero

This theorem can be expressed in terms of linear transformations,

> **Theorem 2.5.2** Key Theorem
>
> Let $T : V \to V$ a linear map, then the following are equivalent:
>
> 1. A is **Invertible**
>
> 2. $\det(A) \neq 0$ where the determinant relies on a choice of basis on V.
>
> 3. $\ker(A) = 0$
>
> 4. If b is a vector in V, then there is a unique vector $v \in V$ satisfying $T(v) = b$
>
> 5. For any basis $\{v_1, \ldots, v_n\}$ of V, $T(v_1), \ldots, T(v_n)$ are linearly independent.
>
> 6. For any basis $\{v_1, \ldots, v_n\}$ of V, if S denotes the transpose linear transformation of T, then $S(v_1), \ldots, S(v_n)$ are linearly independent.
>
> 7. The transpose of T is invertible. (Here the transpose relies on a choice of basis on V)
>
> 8. All the eigenvalues of T are non-zero

The definitions of the determinant and transpose of a linear transformation directly depend on a choice of basis. However, the fact that $\det(T) \neq 0$ does **not** depend on the basis chosen. Linear independence does not rely on choice of basis either.

2.6 Similar Matrices

When choosing a basis for an n-dimensional vector space V, any linear map $T : V \rightarrow V$ can be represented by an $n \times n$ matrix. However, if the choice of basis changes, then the matrix representation of the same linear map will change as well.

Two matrices that represent the same linear transformation under different bases will be **similar**.

> **Definition 2.6.1** Similar Matrices
>
> Two $n \times n$ matrices A, B are **Similar** if there is an invertible matrix C such that
>
> $$A = C^{-1}BC$$

Given two bases for V, β and γ, and any $v \in V$, we need a matrix such that

$$C[v]_\beta = [v]_\gamma$$

The matrix that makes this possible is

$$C = \begin{pmatrix} | & & | \\ [\gamma_1]_\beta & \cdots & [\gamma_n]_\beta \\ | & & | \end{pmatrix}$$

Then if $T : V \rightarrow V$, $A = [T]_\beta$, $B = [T]_\gamma$, and we have $CA = BC$ or $A = C^{-1}BC$, then A and B are similar.

> **Example 2.6.1** (Similar Matrices)
>
> Let us see the same linear map under two different bases.
>
> Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ x \end{pmatrix}$.
>
> Under the usual basis, $E = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$,
>
> $$[f]_E^E = \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}$$
>
> Under the basis $\beta = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$
>
> $$f\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad f\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$
>
> $$\left[f\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]_\beta = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \left[f\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right]_\beta = \left[\begin{pmatrix} 2 \\ 1 \end{pmatrix} \right]_\beta = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$
>
> Then,
>
> $$[f]_\beta^\beta = \begin{pmatrix} 0 & -1 \\ 0 & 2 \end{pmatrix}$$
>
> Our matrix C is:
>
> $$E_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = -\beta_1 + \beta_2, \implies [E_1]_\beta = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$
>
> $$E_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \beta_1 + 0\beta_2, \implies [E_2]_\beta = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
>
> $$C = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}, \qquad C^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

Such that for $a \in \mathbb{R}^2$ $C[a]_E = [a]_\beta$

That way, $[f]_E^E$, $[f]_\beta^\beta$ should be similar.

$$[f]_E^E = \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} = C^{-1}[f]_\beta^\beta C$$

2.7 Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors can be used to pick a better basis for a particular linear transformation.

Definition 2.7.1 Eigenvalues and Eigenvectors

Let $T : V \to V$ be a linear transformation. Then a non-zero vector v will be an **Eigenvector** of T with **Eigenvalue** λ, a scalar, if

$$T(v) = \lambda v$$

For an $n \times n$ matrix A, it is the same idea, non-zero column vector $x \in \mathbb{R}^n$ eigenvector with λ eigenvalue if

$$Ax = \lambda x$$

Geometrically, v is an eigenvector of T if T 'stretches' v by a factor of λ.

Proposition 2.7.1

A scalar λ will be an **eigenvalue** of a square matrix A iff λ is a root of the polynomial

$$P(t) = \det(tI - A)$$

The polynomial $P(t)$ is called the **Characteristic Polynomial**.

Theorem 2.7.1

Let A, B be similar matrices, then their characteristic polynomial is the same.

Corollary 2.7.1

The Eigenvalues for similar matrices are equal.

However, this does **not** imply matrices with the same eigenvalues are similar.

Theorem 2.7.2

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues, counted with multiplicity, of a matrix A. Then

$$\det(A) = \lambda_1 \ldots \lambda_n$$

Counting with multiplicity implies counting roots like 2 in $P(x) = (x-2)^3$ three times.

How good a basis is for a transformation depends on how close it is to being diagonal. This is far from being always possible though, but there are techniques that make is slightly better like the Jordan Normal Form.

These all involve finding eigenvalues, their eigenvectors, or their generalized form.

2.8 Exercises

Let $L : V \to W$ be a linear map between two vector spaces. Show that

$$\dim(\ker(L)) + \dim(\mathrm{Im}(L)) = \dim(V)$$

Solution: Let $\{v_1, \ldots, v_m\}$ be a basis for $\ker(L)$, with $\dim(V) = n$. This basis may not be sufficient to span V, so we add linearly independent vectors until it is a basis of V.

Then, $\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ is a basis for V.

We will show $\{L(v_{m+1}), \ldots, L(v_n)\}$ is a basis for $\mathrm{Im}(L)$.

Let $v \in V$, we know that

$$v = a_1 v_1 + \cdots + a_m v_m + a_{m+1} v_{m+1} + \cdots + a_n v_n$$

Then,

$$\begin{aligned}
L(v) &= L(a_1 v_1 + \cdots + a_m v_m + a_{m+1} v_{m+1} + \cdots + a_n v_n) \\
&= a_1 L(v_1) + \cdots + a_m L(v_m) + a_{m+1} L(v_{m+1}) + \cdots + a_n L(v_n) \\
&= 0 + \cdots + 0 + a_{m+1} L(v_{m+1}) + \cdots + a_n L(v_n)
\end{aligned}$$

$L(v_i)$ is zero if $i = 1, \ldots, m$, because those are in $\ker(L)$. Then, any element in $\mathrm{Im}(L)$ is a linear combination of $L(v_{m+1}), \ldots, L(v_n)$

Suppose there are $a_{m+1}, \ldots a_n$ such that

$$a_{m+1} L(v_{m+1}) + \cdots + a_n L(v_n) = 0$$

Then,

$$L(a_{m+1} v_{m+1} + \cdots + a_n v_n) = 0 \quad \Longrightarrow \quad a_{m+1} v_{m+1} + \cdots + a_n v_n \in \ker(L)$$

Since $\{v_1, \ldots, v_m\}$ spans $\ker(L)$, there are $b_1, \ldots, b_m$ such that

$$\begin{aligned}
a_{m+1} v_{m+1} + \cdots + a_n v_n &= b_1 v_1 + \cdots + b_n v_m \\
-b_1 v_1 - \cdots - b_m v_m + a_{m+1} v_{m+1} + \cdots + a_n v_n &= 0
\end{aligned}$$

Now, since $\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ is a basis for V (then linearly independent), and this linear combination equals to zero, the b_js and, in particular, the a_is should all be zero. This implies that $L(v_{m+1}), \ldots, L(v_n)$ are all linearly independent, and they span $\mathrm{Im}(L)$; hence, they are a basis of $\mathrm{Im}(L)$.

Calculating the dimension,

$$\dim(\ker(L)) + \dim(\mathrm{Im}(L)) = m + n - m = n = \dim(V)$$

☺

Consider the set of all polynomials in one variable with real coefficients with degree less or equal to 3.

(a) Show that this set forms a vector space of dimension four.

(b) Find a basis for this vector space.

(c) Show that differentiating a polynomial is a linear map.

(d) Given the basis in (b), write down the matrix representative of the derivative.

Solution:

(a) Let $P_3 = \{a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_i \in \mathbb{R}\}$. Manually checking the conditions for a vector space here works. The set $\{1, x, x^2, x^3\}$ is a linearly independent set since

$$a_0 + a_1x + a_2x^2 + a_3x^3 = 0$$

does imply $a_i = 0\ \forall i$.

The set also evidently spans P_3 since all polynomials are already a linear combination of these elements. Since the basis has four elements, $\dim(P_3) = 4$.

(b) $\{1, x, x^2, x^3\}$.

(c) Let's consider the map $\frac{d}{dx} : P_3 \to P_2$ defined by

$$\frac{d}{dx}(a_0 + a_1x + a_2x^2 + a_3x^3) = a_1 + 2a_2x + 3a_3x^2$$

Given two polynomials in P_3, $a_0 + a_1x + a_2x^2 + a_3x^3$ and $b_0 + b_1x + b_2x^2 + b_3x^3$

$$\frac{d}{dx}\left(\lambda_1(a_0 + a_1x + a_2x^2 + a_3x^3) + \lambda_2(b_0 + b_1x + b_2x^2 + b_3x^3)\right)$$
$$= \frac{d}{dx}\left((\lambda_1a_0 + \lambda_2b_0) + (\lambda_1a_1 + \lambda_2b_1)x + (\lambda_1a_2 + \lambda_2b_2)x^2 + (\lambda_1a_3 + \lambda_2b_3)x^3\right)$$
$$= (\lambda_1a_1 + \lambda_2b_1) + 2(\lambda_1a_2 + \lambda_2b_2)x + 3(\lambda_1a_3 + \lambda_2b_3)x^2$$
$$= \lambda_1(a_1 + 2a_2x + 3a_3x^2) + \lambda_2(b_1 + 2b_2x + 3b_3x^2)$$
$$= \lambda_1\frac{d}{dx}(a_0 + a_1x + a_2x^2 + a_3x^3) + \lambda_2\frac{d}{dx}(b_0 + b_1x + b_2x^2 + b_3x^3)$$

And $\frac{d}{dx}$ is linear.

☺

(d) Given the bases $\beta = \{1, x, x^2, x^3\}$ and $\gamma = \{1, x, x^2\}$ for P_3 and P_2,

$$\frac{d}{dx}(1) = 0, \quad \frac{d}{dx}(x) = 1, \quad \frac{d}{dx}(x^2) = 2x, \quad \frac{d}{dx}(x^3) = 3x^2,$$

When translating to the basis γ,

$$\left[\frac{d}{dx}(1)\right]_\gamma = \begin{pmatrix}0\\0\\0\end{pmatrix}, \quad \left[\frac{d}{dx}(x)\right]_\gamma = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \left[\frac{d}{dx}(x^2)\right]_\gamma = \begin{pmatrix}0\\2\\0\end{pmatrix}, \quad \left[\frac{d}{dx}(x^3)\right]_\gamma = \begin{pmatrix}0\\0\\3\end{pmatrix},$$

So the matrix representation of $\frac{d}{dx}$ is:

$$\left[\frac{d}{dx}\right]_\beta^\gamma = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Using it on $p(x) = 3 + 7x + 5x^3$ for example,

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ 7 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \\ 15 \end{pmatrix}$$

Which corresponds to $7 + 15x^2$.

I had never seen this way of portraying the derivative as a matrix. This must be the way computers do derivatives, unless there is an even more efficient way to do them, matrix multiplication is very costly.

Question 3

Let $T : V \to W$ be a linear transformation from a vector space V to a vector space W. Show that $\mathrm{Im}(T)$ is a subspace of W.

Solution: $0 \in \mathrm{Im}(T)$, since T is linear, hence $T(0) = 0$.

If $w_1, w_2 \in \mathrm{Im}(T)$, then there exists v_1, v_2 such that $T(v_1) = w_1$, $T(v_2) = w_2$. Then,

$$w_1 + w_2 = T(v_1) + T(v_2) = T(v_1 + v_2) \in \mathrm{Im}(T)$$

Also, if λ is a scalar,

$$\lambda w_1 = \lambda T(v_1) = T(\lambda v_1) \in \mathrm{Im}(T)$$

Due to these criteria $\mathrm{Im}(T)$ is a subspace of W.

☺

Question 4

Let A, B be two invertible matrices. Show that

$$(AB)^{-1} = B^{-1}A^{-1}$$

Solution: Calculating

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \\ &= AIA^{-1} \\ &= AA^{-1} \\ &= I \end{aligned}$$

Now the other way,

$$\begin{aligned} (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B \\ &= B^{-1}IB \\ &= BB^{-1} \\ &= I \end{aligned}$$

Since $(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I$, it is clear that $(AB)^{-1} = B^{-1}A^{-1}$.

☺

Let

$$A = \begin{pmatrix} 2 & -2 \\ 1 & 5 \end{pmatrix}$$

Find a matrix C such that $C^{-1}AC$ is diagonal.

Solution: We first need to find the eigenvalues for the matrix.

$$\det(A - \lambda I) = (2 - \lambda)(5 - \lambda) + 2 = \lambda^2 - 7\lambda + 12 = (\lambda - 3)(\lambda - 4)$$

$$\lambda_1 = 3, \quad \lambda_2 = 4$$

Since $Av = \lambda v$, $Av - \lambda v = 0$, $(Av - \lambda I)v = 0$

Solving for V_1

$$(Av - \lambda_1 I \,|\, 0) = \left(\begin{array}{cc|c} -1 & -2 & 0 \\ 1 & 2 & 0 \end{array} \right) \implies \begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array} \implies v_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

$$(Av - \lambda_2 I \,|\, 0) = \left(\begin{array}{cc|c} -2 & -2 & 0 \\ 1 & 1 & 0 \end{array} \right) \implies \begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array} \implies v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Then,

$$C = \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} -1 & -1 \\ 1 & 2 \end{pmatrix}$$

Calculating $C^{-1}AC$

$$\begin{pmatrix} -1 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -2 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -3 & -3 \\ 4 & 8 \end{pmatrix}\begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$