Linear Algebra

Zac Zerafa

February 11, 2026

Contents

Part I

Fundamentals

Chapter 1

Linear equations

Descartes' analytic geometry was an initial precursor. As the methods of algebra, geometry, and the general formalism of mathematics have advances, the field of linear algebra has come to full fruition. Indeed it is one of the most fundamental, powerful, and versatile tools in the arsenal of a mathemtician, lying at the heart of many mathematic fields relating to geometry and mathematical analysis, mathematical optimization, but also finding applications elsewhere.

Before understanding these methods, it is necessary to understand the nature of linear equations; many calculations in linear algebra will be reduced to these so it is important that we are able to understand the nature of solutions to such problems and how they are calculated.

1.1 Gaussian elimination

Readers are familiar with the idea of simultaneous equations from elementary algebra. The methods of *substitution or elimination* may be used to find a solution to such a system of equations.

We develop a concrete algorithm to solve linear systems of equations based on the idea of elimination.

1.1.1 Non-unique solutions

Chapter 2

Matrix theory

2.1 Matrixes and Vectors

Matrixes and vectors introduce a notation that simplifies systems of linear equations into one 'matrix equation'.

Study on the properties of matrixes completes our understanding about solutions to systems of linear equations, and allows us to develop a deeper theory about them; they will become indispensable tools when we study *linear spaces*.

However before we begin, we must acknowledge the existence of an algebraic structure called a *field*.

We mostly deal with the field $\mathbb{R}$, however the sets $\mathbb{Q}, \mathbb{C}$ also form fields; indeed there are many more fields out there. We don't require deep knowledge about fields (ring theory) to do the linear algebra of this book.

Definition 2.1. A *scalar* is an element of a field

Definition 2.2 (Matrix)**.** A $n \times m$ *matrix* is a grid with n rows (horizontal) and m columns (vertical) that indexes mathematical objects. WHen all its indexes are members of the same field F, we call it a *matrix over the field F*

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \dots & \mathbf{A}_{1m} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{n1} & \dots & \mathbf{A}_{nm} \end{bmatrix}$$

Usually the objects indexed in a matrix are numbers, however they could be anything from functions, random variables, sets, etc. In this book, we will mostly be using matrixes whose entries are elements of $\mathbb{R}$ (matrixes over the field $\mathbb{R}$).

Definition 2.3 (Column vector)**.** A *n-coordinate vector* is a $n \times 1$ matrix.

put in 3.tex Column vectors are often used for finite dimension linear spaces to represent vectors with respect to their standard basis. $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \sum_{k=1}^{n} c_k \mathbf{e}_k$

Definition 2.4 (Square matrix)**.** A *square matrix* a $n \times n$ matrix. In other words,it is a matrix such that the amount of rows equals the amount of columns.

2.2 Matrix algebra

Just like numbers, there are operations defined on matrixes. We can add two matrixes, multiply two matrixes, and scale a matrix by a number In primary school, one learns numbers, and then algorithms to add and multiply them. When one learns 'algebra' in highschool, they learn manipulation laws for solving algebraic equations.

When dealing with matrixes, we will follow a similar program; we will discuss what matrixes are, provide algorithms to define what it means to add and multiply (and scale; more on this later) matrixes, and finally discuss the manipulation laws of 'matrix algebra'.

It is important to remember that in these definitions, we are assuming that entries are elements of $\mathbb{R}$. When matrixes were defined earlier, we allowed the possibility of its entries to contain any type of mathematical object; in this chapter we are defining our matrix algebra only for when entries are elements of $\mathbb{R}$!

Definition 2.5 (Identity matrix)**.** The *identity matrix of n dimensions* is the square matrix $\mathbf{I}_n$ with dimensions $n \times n$ with the following entries. Using matrix indexing notation, we have the folliwing.

$$(\mathbf{I}_n)_{ij} = \begin{cases} 1 & i = j \\ 0 i \neq j \end{cases}$$

When context regarding the dimension of the ideneity matrix is clear, one may write the identity matrix as $\mathbf{I}$.

Definition 2.6 (Zero matrix)**.** A *zero matrix of* $n \times m$ *dimensions* is the $n \times m$ matrix where all entries zero.

2.2.1 Matrix addition and scalar multiplication

Matrix addition works 'component wise'; the entries of the same index of both matrixes is added together to make that index for the resultant matrix. That means only matrixes of the same dimension may be added together.

$$(\mathbf{A} + \mathbf{B})_{ij} = \mathbf{A}_{ij} + \mathbf{B}_{ij}$$

Matrix scalar multiplication works similarly

$$(k\mathbf{A})_{ij} = k\mathbf{A}_{ij}$$

2.2.2 Laws of matrix addition and scalar multiplication

Matrix addition and scalar multiplication is faithful to regular real number addition and multuplication.

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$$

$$\mathbf{A} + \mathbf{0} = \mathbf{B} + \mathbf{A} = \mathbf{A}$$

$$k(\mathbf{A} + \mathbf{B}) = k\mathbf{A} + \mathbf{B} = \mathbf{A}$$

$$\exists \mathbf{A}, \mathbf{B}(\mathbf{AB} \neq \mathbf{BA})$$

$$\forall \mathbf{A}(\exists - \mathbf{A}[\mathbf{A} + (-\mathbf{A} = \mathbf{0}])$$

This last law states that for any matrix, there is some 'inverse matrix' you can add to it to get the zero matrix. This can be used to define matrix subtraction.

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$$

2.2.3 Matrix multiplication

2.2.4 Laws of matrix multiplication

Matrix multiplication unfortunately isn't as liberal as real number multiplication. Matrix multiplication multiplies entries of the first matrix' row by entries of the second matrix's columns, and takes their sum.

$$(\mathbf{AB})_{ij} = \sum_{k=1}^{n} \mathbf{A}_{ik}\mathbf{B}_{kj}$$

- Matrix multiplication by the identity matrix returns the original matrix
- Matrix multiplication doesn't necessarily commute (order of matrixes may change result)

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}$$
$$\exists \mathbf{A}, \mathbf{B}(\mathbf{AB} \neq \mathbf{BA})$$
$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$$

Are there 'inverse matrixes' with respect to matrix multiplication? Yes, however the conditions for this are not so straightfoward in comparison to matrix addition.

In the case of real numbers, the only real number that cannot be multiplied to make 1 is 0. For square matrixes, that cannot be matrix mutliplied by anything to make the identity matrix have a *determinant* of 0; we require new theory to understand this.

2.2.5 Relationship to linear equations

We have developed a whole new type of mathematical object and defined how operations act upon them. Perhaps it is starting to take shape why all these definitions are being made.

A system of linear equations can be reformulated in terms of in terms of matrixes and vectors; we can treat a system of linear equations as a simgle matrix equation.

Matrix multiplication is designed in such a way that given a vector $\mathbf{x}$ filled with variables, a matrix $\mathbf{A}$ filled with coefficients (where the first row is for the first equation's coefficients, second row is for the second equation's coefficients etc.), and a vector of right hand side numbers $\mathbf{b}$, our entire system of linear equations can be written simply as $\mathbf{Ax} = \mathbf{b}$.

What's more is that we can think of the problem as such; given a 'function' $T(\mathbf{x}) = \mathbf{Ax}$, when do we have $T(\mathbf{x}) = \mathbf{b}$? We are familiar with finding the x such that $f(x) = y$; systems of linear equations now have the analogue of finding the $\mathbf{x}$ such that $T(\mathbf{x}) = \mathbf{b}$! This kind of function is called a 'linear map'; we will study these in this book.

In a basic sense, mathematicians can use matrixes like 'arrays'; notational 'boxes' employed as a way to conveniently index elements, but the true power of matrixes comes from their ability to define a unique 'linear map'. Even though at first matrixes just look like boxes that hold numbers, they can also linear functions that map vectors to vectors! Matrix multipliciation can even be thought of as composition of the linear maps induced by these matrixes, but again, we'll savor the details for later.

From now on, we will treat systems of linear equations in terms of matrix theory, that is, we will write systems of linear equations as $\mathbf{Ax} = \mathbf{b}$.

Although we are discussing matrixes

2.3 Determinant

For polynomial functions, the discriminant tells us whether there exist any real solutions to the polynomial. For systems of linear equations (matrixes), the determinant tells us whether a unique solution exists.

Definition 2.7 (Determinant function)**.**

The discriminant and determinant are similar in this aspect, and this is the historical origin of the determinant. However, the determinant has

2.3.1 Invertible matrix

Reciprocating a number is a 'multiplicative inverse', division is essentially multiplying by such an 'inverse element' since $a \cdot \frac{1}{b} = a \div b$. The only number that isn't 'invertible' is 0; $\frac{1}{0}$ is undefined. Can we develop a similar system for matrixes? Is every matrix invertible?

Definition 2.8 (Invertible matrix)**.** An *invertible matrix* is a square matrix $\mathbf{A}$ such that there exists some $\mathbf{A}^{-1}$ such that the following holds.

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

Such an $\mathbf{A}^{-1}$ is called the *inverse matrix of* A

2.3.2 Invertible matrix theorem

Unfortunately, not evey matrix is invertible. $\mathbf{0}$ is a counterexample, however there turn out to be many more; is there a 'test' or condition we can use to check invertibility?

Theorem 2.1 (Invertible matrix theorem)**.**

$$\mathbf{A} \text{ is invertible} \iff \det(\mathbf{A}) \neq 0$$

2.3.3 Properties of the determinant

Proposition 2.1.

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$$
$$\det(\mathbf{A}^\intercal) = \det(\mathbf{A})$$
$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$$
$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$$

2.3.4 Laplace expansion

2.3.5 Leibniz' formula

2.3.6 Cramer's rule

2.3.7 Principle minors

2.3.8 Matrix determinant lemma

2.4 Transposition

Definition 2.9 (Transposition)**.**

$$(\mathbf{A}^\intercal)_{ij} = (\mathbf{A})_{ji}$$

Definition 2.10 (Conjugate transposition)**.**

$$(\mathbf{A}^*)_{ij} = (\mathbf{A}^\intercal)^*_{ij}$$

2.4.1 Symmetric matrix

Definition 2.11 (Symmetric matrix)**.** A *symmetric matrix* is a square matrix such that it is equal to its transpose.

$$\mathbf{A} \text{ is symmetric} \iff \mathbf{A} = \mathbf{A}^\intercal$$

Definition 2.12. A *positive-definite matrix*

$$\mathbf{A} \text{ is positive-definite} \iff \forall \mathbf{x}[\mathbf{x}^\intercal \mathbf{A}\mathbf{x} > 0]$$

A *negative-definite matrix*

$$\mathbf{A} \text{ is negative-definite} \iff \forall \mathbf{x}[\mathbf{x}^\intercal \mathbf{A}\mathbf{x} < 0]$$

Definition 2.13. A *positive semi-definite matrix* A *negative semi-definite matrix*

2.4.2 Orthogonal matrix*

We introduce it for the sake of completeness; we will explore why such a matrix is meaningful once more linear algebra is understood.

Definition 2.14 (Orthonormal matrix)**.**

$$\mathbf{A}^{-1} = \mathbf{A}^\intercal$$

2.4.3 Hermitian matrix*

Definition 2.15 (Hermitian matrix)**.**

$$\mathbf{A} = (\mathbf{A}^\intercal)^*$$

- hermitian matrix

Definition 2.16 (Unitary matrix)**.**

$$\mathbf{A}^* = \mathbf{A}^{-1}$$

2.5 Miscellaneous formulae

trace Sherman-Morrison formula

Chapter 3

Linear spaces

Much of the time, students have previous experience with vectors in the form of *Euclidean vectors* and even representing them as *coordinate vectors* as defined in the previous chapter.

Linear spaces are an algebraic structure that allows a general framework to define and handle vectors.

We would like a system that defines vectors based on their algebraic behaviour than being confined to a single, concrete coordinate implementation; this could help us translate problems stated in different coordinate implementations with more ease.

This would also allow mathematicians to abstract the algebraic behaviour of vectors and study mathematical objects with the same properties. For example, functional analysis makes use of linear algebra by studying functions as vectors (rather studying function spaces as linear spaces)!

This is because a linear space is an algebraic structure; it doesn't require a literal interpretation of 'direction' to be; it just needs a set of our synthetically defined objects (that we will eventually call vectors) to have certain algebraic properties.

3.1 Linear space

Behold, the ultimate algebraic structure at our disposal to categorize the behaviour of vectors.

Definition 3.1 (Linear space)**.** A *linear space* is an ordered pair $(V, F, +, \cdot)$ of a field and a set of objects called *vectors.* We say that V is a linear space over F.

- $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- $\exists \mathbf{0} \in V[\forall \mathbf{v} \in V[\mathbf{v} + \mathbf{0} = \mathbf{v}]]$
- $\forall \mathbf{v} \in V[\exists (-\mathbf{v}) \in V[\mathbf{v} + (-\mathbf{v}) = \mathbf{0}]]$
- $c(d\mathbf{v}) = (cd)\mathbf{v}$
- $c(d\mathbf{v}) = (cd)\mathbf{v}$
- $1_F \mathbf{v} = \mathbf{v}$
- $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
- $(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}$

The mention of a field is a notion that readers may not be familiar with.

A field is a type of algebraic structure of a set where some notion of 'addition' and 'multiplication' are defined and follow some other special properties (in many cases, these *are* the addition and multiplication that we are familiar with).

For most situations we will come across, F is the set of real or complex numbers.

Linear spaces are actually a special clas of algebraic structures called *modules*, however often one learns linear algebra before ring theory, so we accept the completely equivalent and more accessible definition above.

3.1.1 Properties of linear spaces

$$0_F \mathbf{v} = \mathbf{0}$$
$$\exists! \mathbf{0} \in V[\forall \mathbf{v} \in V[\mathbf{v} + \mathbf{0} = \mathbf{v}]]$$
$$\forall \mathbf{v} \in V[\exists! (-\mathbf{v}) \in V[\mathbf{v} + (-\mathbf{v}) = \mathbf{0}]]$$
$$c\mathbf{0} = \mathbf{0}$$
$$(-c\mathbf{v}) = c(-\mathbf{v})$$
$$c\mathbf{v} = \mathbf{0} \implies c = 0_F \vee \mathbf{v} = \mathbf{0}$$

We will define subtraction of vectors in the following manner.

$$\mathbf{v} - \mathbf{u} = \mathbf{v} + (-\mathbf{u})$$

3.1.2 Examples of linear spaces

Example 3.1 (Real coordinate spaces). $(\mathbb{R}^n, \mathbb{R}, +, \cdot)$

This is by far the most common type of linear space that one encounters, and you have probably used it implicity numerous times! It is the cornerstone for mathematical analysis, cartesian geometry, it is an object of particular interest in topology, and outside pure mathematics it is fundamental in physics; it's influence resounds throughout mathematics as a whole.

Example 3.2 (Linear space of polynomials up to degree 5).

Notice that due to the requirements of a linear space, this space includes the zero polynomial, has dimension 6. Such a linear space may be useful in the area of field theory.

Example 3.3 (Linear spaces of solutions of linear ODEs).

Since solutions of linear ODEs form a linear space, linear algebra becomes an extremely powerful tool in the analysis of solutions, for instance by the use of the Wronskian.

3.2 Linear subspace

Perhaps only a handful of the vectors within V are enough to satisfy the properties of a linear space. When we consider subsets of linear spaces that form their own linear space, we call this a *linear subspace*.

Definition 3.2 (Linear subpsace). A *linear supspace* of $(V, F, +, \cdot)$ is a linear space $(W, F, +, \cdot)$ such that W is a subset of V. $W \leq V$ denotes that W is a subgroup of V.

$$W \leq V \iff (W \subseteq V) \land (W, F, +, \cdot) \text{ is a linear space}$$

3.2.1 Examples of linear subspaces

Proposition 3.1 (Intersection subspace).

$$(V, F, +, \cdot)$$
$$(W, F, +, \cdot)$$
$$V \cap W \leq V$$
$$V \cap W \leq W$$

Proposition 3.2 (Additive superspace)**.**

$$(V, F, +, \cdot)$$

$$(W, F, +, \cdot)$$

$$V \leq V + W$$

$$W \leq V + W$$

When we eventually study linear maps, we will see that they add substantial contributions to the theory of subspaces.

3.3 Linear combinations

Studying linear combinations is fundamental for analyzing linear spaces since it is a precursor to the notion of a *basis*, which is perhaps the ultimate tool for studying a linear space.

Definition 3.3 ((finite) Linear combination)**.** For a set of vectors and set of coefficients, a *linear combination* is a finite sum of vectors scaled by some scalar.

$$\sum_{i=1}^{n} c_i \mathbf{v}$$

3.3.1 Linear span

To represent the set of all linear combinations possible with a certain set of vectors, we use a *linear span*

Definition 3.4 (Linear span)**.** (F, V)

$$S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$$

$$\text{span}(S) = \{\sum_{i=1}^{n} c_i \mathbf{v}_i : c_i \in F\}$$

$$A \subseteq V \wedge |A| < \aleph_0$$

$$\text{span}(A) = \{\sum_{\mathbf{v} \in B} c_{\mathbf{v}} \mathbf{v} : c_{\mathbf{v}} \in F \wedge B \subseteq A \wedge |B| < \aleph_0\}$$

Proposition 3.3. Linear spans form a subspace.

3.3.2 Linear independence

Representing vectors by means of linear combinations is nice, however what's even better would be if vectors have a *unique* linear combination, given that we're using some fixed set of spanning vectors.

Definition 3.5 (Linearly independent set)**.** A set of vectors A is *linearly independent set* iff no element $\mathbf{v} \in A$ can be represented as a linear combination of element sin $A \setminus \{\mathbf{v}\}$

$$A \text{ is linearly independent} \iff [\nexists \mathbf{u} \in A[\exists B \subseteq A[|B| < \aleph_0 \wedge \sum_{\mathbf{v} \in B} c_{\mathbf{v}}\mathbf{v} = \mathbf{u} \wedge c_{\mathbf{v}} \in F]]$$

There are alternative equivalent ways that a linearly independent set may be defined; different definitions give more intuition for the meaning of the concept, plus they may facilitate constructing a proof.

Proposition 3.4 (Equivalent definitions for linearly independent set)**.** •

no element $\mathbf{v} \in A$ can be represented as a linear combination of element sin $A \setminus \{\mathbf{v}\}$

- There is no nontrivial linear combination of vectors in A for $\mathbf{0}$
- The determinant of the matrix whose columns are the vectors in A is 0

$$A \text{ is linearly independent} \iff \nexists B \subseteq A[|B| < \aleph_0 \wedge \sum_{\mathbf{v} \in B} c_{\mathbf{v}}\mathbf{v} = \mathbf{0} \wedge c_{\mathbf{v}} \in F]$$

3.4 Basis (Linear space)

Definition 3.6 (Basis of a linear space)**.** A *basis of* V is a set of linearly independent vectors B such that every element in V is a linear combination of the vectors in B. $|B|$ is known as the *rank of* B.

$$\mathcal{B} \text{ is a basis of } V \iff \mathcal{B} \text{ is linearly independent } land \text{span}(\mathcal{B}) = V$$

Definition 3.7.

Definition 3.8 (Basis of a linear space)**.** A *basis of* V or *Hamel basis of* V is a set of linearly independent vectors B such that every element in V is a linear combination of the vectors in B. $|B|$ is known as the *rank of* B.

$$\mathcal{B} \text{ is a basis of } V \iff \mathcal{B} \text{ is linearly independent } land \text{span}(\mathcal{B}) = V$$

Definition 3.9. The term 'Hamel basis' is more common in the stuy of functional analysis where (spoiler alert), there are different types of basis' needed to generate more complicated spaces; we will not consider such things in this book.

Proposition 3.5 (Equivalent definitions of a basis)**.**

- $\text{span}(\mathcal{B}) = V$ and the vectors of $\mathcal{B}$ are linearly independent to eachother
- There is no nontrivial linear combination for $\mathbf{0}$

3.5 Dimension

The notion of a basis is extremely useful, however can we *always* analyze a linear space by means of a basis? This is an interesting question because it relies on the foundation of mathematics that we use.

We'll start by considering this over *finite linear spaces*.

Definition 3.10 (Finite linear space)**.** A finite linear space is a linear space spanned by a set $AsubseteqV$ of finite cardinality.

$$V \text{ is a finite linear space} \iff \exists A \subseteq V[|A| < \aleph_0 \wedge \text{span}(A) = V]$$

Assuming we have a finite linear space, we can prove this with no qualms.

Proposition 3.6. Let $(V, F, +, \cdot)$ be a finite linear space, there exists some $\mathcal{B} \subseteq V$ such that $\mathcal{B}$ is a basis of V.

We require the use of an axiom that was once controversial among mathematicians; the axiom of choice. As it turns out, infinite linear spaces always having a basis is actually logically equivalent to the axiom of choice, making the following result not just of interest in linear algebra but also in mathematical logic.

Proposition 3.7. Let $(V, F, +, \cdot)$ be a infinite linear space, there exists some $\mathcal{B} \subseteq V$ such that $\mathcal{B}$ is a basis of V iff we accept the axiom of choice.

In this book we accept the axiom of choice, so we accept the previous proposition. That said, most work (especially in the 'fundamentals' part of the book) focuses on results pertaining to finite linear spaces.

Now we know that all linear spaces have a basis, but what is even nicer is the following fact.

Proposition 3.8. Let $(V, F, +, \cdot)$ be a linear space, all its basis' have the same cardinality.

Since all basis' for the same linear space must have the same cardinality, we can define *dimension* as an intrinsic property of a linear space pertaining to the cardinality of any basis it can produce.

Definition 3.11 (Dimension of a linear space)**.** Let $(V, F, +, \cdot)$ be a linear space, the *dimension of* V is the cardinality of any basis of V. We write $\dim(V)$ to represent this cardinality.

Corollary 3.1. For $(V, F, +, \cdot)$ Any set of more than $\dim(V)$ vectors is linearly independent.

Our proof of the existence of a basis for any lienar space relies on a method of 'cherry picking' linearly independent vectors until a basis is constructed; this proof is nice because it is constructive! The proof essentially corresponds to an algorithm on how to generate a basis. For the finite linear spaces, we only need to pick linearly independent vectors until we reach the dimension.

Theorem 3.1 (Basis theorem)**.** Let $(V, F, +, \cdot)$ be linear space with $\dim(V) = n$, any set of n linearly independent vectors of V is a basis for (F, V)

Proposition 3.9. All finite dimensional linear spaces over F with the same dimension are isomorphic.

$$(V, F, +_V, \cdot_V)$$

$$(W, F, +_W, \cdot_W)$$

$$\dim(V) < \infty \implies [\dim(V) = \dim(W) \iff V \cong W]$$

Though we can construct such an isomorphism for finite dimensional linear spaces, this result doesn't hold for linear spaces with infinite dimension.

This theorem is enlightening because it implies the behaviour of linear spaces is independent of any basis, with the isomorphism itsem being the 'change of basis' function!

A change of basis is often thought of a change of coordinate systems, which means a linear isomorphism onto the same linear space. This is exactly the definition of a *linear endomorphism.*

3.6 Spanning set theorem

Theorem 3.2 (Spanning set theorem)**.** For any $\mathbf{v} \in A$ that is linearly dependent to $A \setminus \{\mathbf{v}\}$, we have $\text{span}(A) = \text{span}(A \setminus \ \mathbf{v})$

3.7 Linear space constructions

3.7.1 Direct sum of linear spaces

Definition 3.12 (Direct sum of linear spaces)**.** Let U, W be linear subspaces of V, then the direct sum $U \oplus W$ contains all vectors that have a unique representation as the sum of a vector in U and vector in W

$$U \leq V$$

$$W \leq V$$

$$U \oplus W \iff \exists! \forall \mathbf{v} \in U \oplus W (\mathbf{u} \in U, \mathbf{w} \in W [\mathbf{v} = \mathbf{u} + \mathbf{w}])$$

One way to think of the direct sum is to think of the disjoint union of the subspaces as well as the zero element $[(U \cup W) \setminus (U \cap W)] \cup \{\mathbf{0}\}$ and then closing it up with vector addition.

3.7.2 Direct product of linear spaces

NOTE; THIS SECTION REQUIRES MORE EXPLANATION

Let F be a field, then F^n is a linear space over F.

Proposition 3.10. direct product and direct sum have a linear isomorphism under finite applications.

3.7.3 Quotient linear space

It is highly reccomended that readers be familiar with group theory before understanding quotient spaces.

Definition 3.13. Left coset (Linear space) Let $W \leq V$, then $\mathbf{v} + W = \{\mathbf{v} + \mathbf{w} : \mathbf{w} \in W\}$

Proposition 3.11. Left coset (Linear space) Left cosets are either equal or disjoint. They form an equivalence relation.

As it turns out, left cosets form their own linear space!

Definition 3.14 (Quotient linear space)**.** A *quotient linear space* $(F, V/U)$ where $U \leq V$ is a linear space over F with the distinct sets $\mathbf{v} + U = \{\mathbf{v} + \mathbf{u} : \mathbf{u} \in U\}$ as its vectors.

Chapter 4

Linear maps

Though the notion of a basis is arguably the most important tool to analyze linear spaces, linear maps are definitely a contender.

These types of functions a powerful tools used to study properties of linear spaces since they 'preserve' the algebraic structure of the domain elements in the codomain space (homomorphism).

Additionally they are also worthy of study in and of themselves, displaying interesting properties as well as finding various practical applications.

Definition 4.1. Let V, W be vector spaces over K, a *linear map* or *linear transform* is a function between linear spaces $T : V \to W$ with the following properties.

- $T(\mathbf{u} +_V \mathbf{v}) = T(\mathbf{u}) +_W T(\mathbf{v})$ (T is additive)
- $T(c\mathbf{v}) = cT(\mathbf{v})$ (T is homogeneous of degree 1)

The set $\mathcal{L}(V, U)$ represents the set of all linear maps $T : V \to U$

The following theorem explains why matrixes are all the rage in linear algebra, and it is because of their intimate relationship to linear maps.

Theorem 4.1. For any linear map $T : \mathbb{F}^n \to \mathbb{F}^m$, there exists a matrix over F ($\mathbf{A}$) such that $T(\mathbf{x}) = \mathbf{Ax}$. Furthermore, any matrix over F defines a linear map.

This theorem gets to the heart of matrixes in the context of linear algebra and why matrix addition and multiplication are defined the way they are.

In these contexts, matrixes are not just arrays of scalars, but are used to characterize and specify linear maps of the form $T : F^n \to F^m$!

Matrix multiplication was defined to reflect the composition of linear maps, matrix addition to reflect the addition of linear maps; solving a simultaneous linear equations is essentially finding what domain elements of these linear maps are mapped to some image element. This allows us to view familiar problems from a more functional perspective; a milestone in our mathematical maturity.

By considering the fact that any finite linear space is isomorphic to F^n , we have the following as a corollary.

Proposition 4.1. If $\dim(\mathrm{dom}(T))$ is finite, then there exists some matrix $\mathbf{A}$ such that $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$.

This reiterates the point that although matrixes can be used as a means of indexing elements in a grid-like fashion, their real power lies in the fact that under matrix multiplication they represent linear maps of finite dimension

Linear transforms on scalars are quite mundane; they merely scale the number by some fixed factor. For larger dimensions however, linear transforms can be interpreted with more geometrical flavour; linear transforms rotate, shear, and scale vectors!

4.0.1 Kernel (linear maps)

Definition 4.2 (Kernel (Linear map))**.** Let V, W be linear spaces over K and $T : V \to W$ a linear map.

$$\ker(T) = \{\mathbf{x} : T(\mathbf{x}) = \mathbf{0}\}$$

4.0.2 Image (linear maps)

Definition 4.3 (Image (Linear map))**.** Let V, W be linear spaces over K and $T : V \to W$ a linear map.

$$\mathrm{Im}(T) = \{T(\mathbf{x}) : \mathbf{x} \in \mathrm{dom}(T)\}$$

4.0.3 Exaples of linear maps

Identity map Shearing maps Rotation maps Dilation maps

$$T(\mathbf{x}) = (c\mathbf{I})\mathbf{x}$$

Reflection map

4.0.4 Linear isomorphisms

Definition 4.4 (Linear isomorphism)**.** Let $(V, F, +, \cdot)$ and $(U, F, +, \cdot)$ be two linear spaces over F. A *linear isomorphism* $f : V \to U$ is a bijective linear map. If there exists an isomorphism between V and U, then the linear spaces are *isomorphic* to eachother, also written as $V \cong U$.

4.0.5 Linear endomorphisms

Definition 4.5 (Linear endomorphism)**.** linear isomorphism between linear spaces over the same set of vectors; When two linear spaces are equal

Proposition 4.2. Let $(V, F, +, \cdot)$ be a linear space, then the set of linear endomorphims where $(f + g)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v}), (cf)(\mathbf{v}) = cf(\mathbf{v})$ is a linear space over F with dimension $\dim^2(V)$

4.0.6 Linear forms

Definition 4.6 (Linear form)**.** Let $(V, F, +, \cdot)$ be a linear space, then a *linear form* is a linear map $f : V \to F$ Mapping the linear space to the field it is over.

Proposition 4.3.

$$T : V \to W \text{ is a linear map} \implies \ker(T) \leq V$$

- Column space - Row space - Rank

$$\operatorname{rank}(\mathbf{A}) = \dim(\operatorname{col}(\mathbf{A})) = \dim(\operatorname{row}(\mathbf{A}))$$

$$\operatorname{nullity}(\mathbf{A}) = \dim(\ker(\mathbf{A}))$$

4.1 Rank-nullity theorem

Theorem 4.2 (Rank-Nullity theorem (Linear maps))**.** Let $\mathbf{T} : V \to W$ be a linear map, then the following holds.

$$\dim(V) = \operatorname{nullity}(T) + \operatorname{rank}(T)$$

Theorem 4.3 (Rank-Nullity theorem (Matrix))**.** Let $\mathbf{A}$ be an $m \times n$ matrix of elements of the field F, then the following holds.

$$n = \mathrm{nullity}(\mathbf{A}) + \mathrm{rank}(\mathbf{A})$$

Chapter 5

Normed linear spaces

NOTE; MOVE THIS CHAPTER AFTER ORTHOGONALITY

We have studied linear spaces as algebraic structures, however we can equip them with certain functions that link linear spaces to applications in topology and analysis.

We will not apply the theory here to other branches of mathematics, however we will discuss some elementary consequences of introducing these functions to a linear space.

One of the basic functions we can consider is the *norm*.

5.1 Norm

The idea of the norm is to define a notion of distance of vectors from the origin.

Definition 5.1 (Norm)**.** Given a vector space $(V, F, +, \cdot)$ where F is the field $\mathbb{R}, \mathbb{C}$, a *norm over* V is a function $\|\cdot\| : V \to \mathbb{R}$ with the following properties.

- $\|\mathbf{v} + \mathbf{u}\| \leq \|\mathbf{v}\| + \|\mathbf{u}\|$
- $\|\lambda\mathbf{v}\| = |\lambda|\|\mathbf{v}\|$
- $\|\lambda\mathbf{v}\| = 0 \implies \|\mathbf{v} = \mathbf{0}$

Definition 5.2 (Normed linear space)**.** A *normed linear space* is an ordered pair $(V, \|\cdot\|)$ of a linear space V over $\mathbb{R}, \mathbb{C}$ and a norm over V.

One can prove a few immediate properties from this definition.

$$\|\|\mathbf{v}\| - \|\mathbf{u}\|\| \leq \|\mathbf{v} - \mathbf{u}\|$$
$$\|\sum_{k=1}^{n} \mathbf{v}_k\| \leq \sum_{k=1}^{n} \|\mathbf{v}_k\|$$
$$\|\lambda\mathbf{v}\| = 0 \iff \|\mathbf{v} = \mathbf{0}$$

Perhaps one of the main motivations for a norm is that it allows the creation of a *distance function*; a function that defines the distance between any 2 vectors. We always want distance functions to obey these properties to match our intuition, and indeed our distance function obeys them all. -triangle ineq - symmetry - —x—=0 iff x=0

In fact, this distance function will actually have some even nicer properties that not all distance functions may have. -trasnlation invariance -scaling invariance

A set with a distance function define upon it is an important object in topology called a *metric space*.

Proposition 5.1. Normed linear spaces have the distance function $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|$. Moreover, this distance function is translation and scaling invariant.

5.1.1 Examples of normed linear spaces

5.2 Inner product spaces

Linear spaces endowned with a special function called an inner product.

5.2.1 Inner product

Definition 5.3 (Inner product)**.** Given a vector space $(V, F, +, \cdot)$ where F is the field $\mathbb{R}, \mathbb{C}$, an *inner product over* V is a function $\langle\cdot,\cdot\rangle : V \times V \to F$ with the following properties.

- $\langle\mathbf{u}, \mathbf{v}\rangle = \langle\mathbf{v}, \mathbf{u}\rangle^*$
- $\langle\lambda\mathbf{u} + \mu\mathbf{w}, \mathbf{v}\rangle = \lambda\langle\mathbf{u}, \mathbf{v}\rangle + \mu\langle\mathbf{w}, \mathbf{v}\rangle$
- $\mathbf{v} \neq \mathbf{0} \implies \langle\mathbf{v}, \mathbf{v}\rangle > 0$

Note that conjugate symmetry implies $\langle \mathbf{v}, \mathbf{v} \rangle$ is a real number, so we can discuss it with respect to the standard partial order over the real numbers ($<$).

Definition 5.4 (Inner product space)**.** An *inner product space* is an ordered pair $(V, \langle \cdot, \cdot \rangle)$ of a linear space V over $\mathbb{R}, \mathbb{C}$ and an inner product over V.

One can prove a few immediate properties from this definition.

$$\langle \mathbf{v}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle = 0$$

$$\langle \mathbf{v}, \mathbf{v} \rangle \in [0, \infty)$$

$$\langle \mathbf{v}, \mathbf{v} \rangle = 0 \iff \mathbf{v} = \mathbf{0}$$

$$\langle \mathbf{v}, \lambda \mathbf{u} + \mu \mathbf{w} \rangle = \lambda^* \langle \mathbf{v}, \mathbf{u} \rangle + \mu^* \langle \mathbf{v}, \mathbf{w} \rangle$$

5.2.2 Canonically induced norm

If we have an inner product space, we can actually use our inner product to create a norm function.

Proposition 5.2.

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

This means that any inner product space is also a normed linear space.

5.2.3 Cauchy-Schwarz inequality

From inner product spaces arises perhaps the most renown inequalities in all of mathematics.

Theorem 5.1 (Cauchy-Schwarz inequality)**.**

$$\langle \mathbf{u}, \mathbf{v} \rangle^2 \leq \langle \mathbf{u}, \mathbf{u} \rangle \cdot \langle \mathbf{v}, \mathbf{v} \rangle$$

5.2.4 Pythagorean theorem

Theorem 5.2 (Pythagorean theorem (Linear algebra))**.**

$$\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u} + \mathbf{v}\|^2 \iff \mathbf{u} \perp \mathbf{v}$$

5.3 Euclidean space

Definition 5.5. A *Euclidean space* is an inner product space $(\mathbf{R}^n, \cdot)$.

- $\mathbb{R}^n$ is the set of n-valued real vectors
- $\cdot$ is the inner product on the Euclidean space, called a *dot product*

5.3.1 Dot product

We have an understanding of perpendicularity; vectors that have directions that are completely

Definition 5.6 (Dot product)**.** Consider the Euclidean space $\mathbb{R}^n$, the *dot product* is a map $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ that follows the following algebraic properties.

- $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
- $(c\mathbf{u} + d\mathbf{w}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) + d(\mathbf{w} \cdot \mathbf{v})$
- $\mathbf{u} \cdot \mathbf{u} \geq 0$

One may recognize that this is essentially the definition of an inner product of $\mathbb{R}^n$ (conjugate symmetry implies symmetry when dealing only with real numbers); It turns out that there is only 1 such function that is an inner product on $\mathbb{R}^n$ so our definition is well defined.

Proposition 5.3 (Alternative definition for dot product)**.** $\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} \mathbf{u}_i \mathbf{v}_i$

Proposition 5.4 (Geometric interpretation for dot product)**.** $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos(\theta)$

Notice that cosine of $\frac{\pi}{2}$ is 0, so the dot product ; this gives us the edge we need to formally define what perpendicularity means in Euclidean space!

Definition 5.7 (perpendicular pair of vectors)**.** In $\mathbb{R}^n$, for any $\mathbf{u}$ and $\mathbf{v}$ in the sert are a *perpendicular pair of vectors* iff $\mathbf{u} \cdot \mathbf{v} = 0$

$$\mathbf{v} \perp \mathbf{u} \iff \mathbf{v} \cdot \mathbf{u} = 0$$

Definition 5.8 (Perpendicular set)**.** A *perpendicular set* is a set of vectors S such that any distinct such that any distinct vectors $\mathbf{v}, \mathbf{u} \in S$ are perpendicular to eachother.

5.3.2 Results of Euclidean geometry

Euclidean geometry is characterized in these spaces, so all the familiar theorems of Euclidean geometry may be expressed in terms of this inner product space. For instance, here is a familiar friend of ours who has had a linear algebra 'glow-up'.

Definition 5.9 (Pythagorean theorem (Linear algebra)).

$$\mathbf{u} \propto \mathbf{v} \implies \|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

Indeed one of the powers of linear algebra is providing a strong foundations for analytic geometry (geometry using coordinate system).

5.3.3 Cross product

Given any vector in $\mathbb{R}^2$, we can find exactly 2 perpendicular vector directions by making use of our knowledge of linear maps to rotate our vector by $\frac{\pi}{2}$; multiplying the vector by the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ does the trick. The other direction is given by the negative of this vector.

If we consider $\mathbb{R}^3$, a given vector has an infinite amount of vector directions perpendicular to it. However given 2 vectors, there are exactly 2 vector directions perpendicular to both vectors. This begs the question; can we define an operator that calculates 1 of the 2 vectors perpendicular to a pair of vectors? The answer is yes, if we imagine some algebraic conditions that such an operator would follow, we end up with the cross product.

Definition 5.10 (Cross product). Consider the Euclidean space $\mathbb{R}^3$, the *cross product* is a map $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ that is linear in both arguments (bilinear map) and follows the following algebraic properties.

- $\mathbf{v} \times \mathbf{v} = \mathbf{0}$
- $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$
- $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$
- $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$

FOr our sanity, we should verify that the definition of the cross product is well defined (i.e this definition defines a unique operator, that is any two operators following these properties are actually the same operator).

From a mathematical point of view, this example is interesting as it takes two arguments and is linear in both; the term for this is a bilinear map. In fact, this particular bilinear map is generalized by the notion of a Grassman algebra; this is covered in the advanced part of this book.

Proposition 5.5 (Alternative definition for cross product)**.** $\mathbf{u}\times\mathbf{v} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{vmatrix}$

Proposition 5.6 (Geometric interpretation for cross product)**.** $\mathbf{u} \times \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \sin(\theta)\hat{\mathbf{n}}$

5.4 Dual spaces

Definition 5.11 (Dual space)**.** Let $(V, F, +, \cdot)$ be a linear space over F, its *dual space* $(V^*, F, +, \cdot)$ is the linear space over F formed by $V^* = \mathcal{L}(V, F)$ (all the linear forms on V)

As usual, we verify that the dual space is indeed a linear space over F otherwise this is ill-defined.

Proposition 5.7. Let $(V, F, +, \cdot)$ be a linear space spanned by n linearly independent vectors $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. Then a basis for V^* are the linear forms $f_1, \ldots, f_n$ that obey the following.

$$f_i(\mathbf{b}_i) = \delta_{ij}$$

Proposition 5.8. Let $(V, F, +, \cdot)$ be a linear space, then its dual space has the same dimension as V.

$$\dim(V^*) = \dim(V)$$

5.4.1 Annihilators

5.4.2 Transpose of linear map

Definition 5.12 (Transpose linear map)**.** Let $(V, F, +, \cdot), (W, F, +, \cdot)$ be a linear spaces over F and $f : V \to W$ be a linear map, then the *transpose of* f is the linear map $f^\intercal : W^* \to V^*$ defined as the following. $f^\intercal(l) = l \circ f$.

Chapter 6

Orthogonality

6.1 Spanning set theorem

6.2 Orthogonal pair of vectors

Definition 6.1 (Orthogonal pair of vectors)**.**

$$\mathbf{v} \perp \mathbf{u} \iff \langle \mathbf{v}, \mathbf{u} \rangle = 0$$

Definition 6.2 (Orthogonal set)**.** An *orthogonal set* is a of vectors that are all orthogonal to eachother.

Definition 6.3 (Orthogonal basis)**.** An *orthogonal basis for V* is a basis for V that is an orthogonal set.

An even more standardized case of orthogonality is *orthonormality*, where all vectors are orthogonal to eachother and have unit norm.

6.3 Orthonormal set

Definition 6.4 (Orthogonal set)**.** An *orthonormal set* is an orthogonal set where all vectors have norm 1.

Definition 6.5 (Orthonormal basis)**.** An *orthonormal basis for V* is a basis for V that is an orthonormal set.

6.4 Orthogonal complement

Proposition 6.1.

$$U \leq V \wedge \dim(V) < \infty \implies V = U \oplus U^{\perp}$$

6.5 Best approximation theorem

6.6 Orthogonal decomposition theorem

6.7 Vector projection on 1 dimensional linear subspace

$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}$$

6.8 Vector projection on n dimensional linear subspace

$$W \leq V \wedge \mathcal{U} \text{ is an orthogonal set } \wedge \mathrm{span}(\mathcal{U}) = W \implies \mathrm{proj}_W(\mathbf{v}) = \sum_{\mathbf{u} \in \mathcal{U}} \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}$$

6.9 Gram-Schmidt process

In practice, one may want to orthogonalize their basis; find an orthonormal basis that spans the same space. There exists an algorithm that constructs such a basis called the *Gram-Schmidt process*. It works by projecting each basis vector onto the space spanned by the current orthogonal set; the result of this identifies the component of the basis element that isn't orthogonal to the set and is subsequenty minused from the vector to form a set orthonormal with the current basis, while still spanning the same space.

Theorem 6.1 (Correctness of Gram-Schmidt process)**.**

$$\mathbf{u}_k = \mathbf{b}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{u}_j}(\mathbf{b}_k)$$

$$\mathrm{span}(\mathcal{B}) = \mathrm{span}(\mathcal{U})$$
$$\mathcal{U} \text{ is an orthogonal set}$$

6.10 Orthogonal matrix

Chapter 7

Eigenequations

7.1 Eigenequations

We have seen that linear transforms exhibit interesting behaviours. Sometimes it is useful to know the domain elements of a linear transform such that the linear transform is merely scaling the vector; this is the idea that gave birth to *eigenequations*, and there will be an abundance of things to say about them mathemtically too.

Definition 7.1. An *eigenequation* is an equation of the following form.

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

$$T(\mathbf{x}) = \lambda\mathbf{x}$$

- $\mathbf{A}$ is a square matrix
- $T : V \to V$ is a linear endomorphism
- solutions of λ are called *eigenvalues*
- solutions of $\mathbf{x}$ are called *eigenvectors*

Let V be a linear space over $\mathbb{C}$, then all eigenequations have nontrivial solutions.

7.1.1 Solving eigenequations

Theorem 7.1 (Characteristic polynomial (Linear algebra))**.**

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \iff \det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

7.1.2 Examples of eigenequations

7.1.3 Eigenvalues

multiplicity of eigenvalues

Eigenvalues of multiplicity n are therefore associated with n eigenvectors.

multiplicity of symmetric eigenequation eigenvalues of positive-definite eigenequation eigenvalues of orthogonal eigenequation Eigenvalues and trace Eigenvalues and determinant

The theory of eigenvalues is closely related to the propeorties of matrix 'definiteness'.

Proposition 7.1. All eigenvalues of a square positive definite matrix are positive.

Proposition 7.2. All eigenvalues of a square positive definite matrix are positive.

7.1.4 Eigenvectors

Theorem 7.2. Let $\mathbf{A}$ be a symmetric matrix. Then its eigenvectors that are associated to different eigenvalues are orthogonal.

7.2 Eigenspaces

spectrum

7.2.1 Eigenspaces

When an eigenvalue has multiplicity greater than 1, its related eigenvectors span an *eigenspace*.

eigenspace

Definition 7.2. An *eigenspace* is a linear space spanned by the eigenvectors associated with the same eigenvalue. These eigenvectors form an *eigenbasis* for this eigenspace.

$$E(\lambda, T) = \{\mathbf{x} : T(\mathbf{x} = \lambda\mathbf{x}\}$$

It is indeed a space because one can easily construct a linear map whose kernel is the same set, and kernels are linear subspaces.

Dimension of eigenspace is multiplicity of its eigenvalue

7.2.2 Properties of eigenspaces

Allusion to use in functional analysis.

7.3 Eigendecomposition and matrix diagonalization

7.3.1 Matrix diagonalization

Definition 7.3 (Diagonalizable matrix)**.** An *diagonalizable matrix* is a square matrix $\mathbf{A}$ such that there exists matrixes $\mathbf{P}, \mathbf{D}$ such that $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$

- $\mathbf{P}$ is an invertible matrix
- $\mathbf{D}$ is a diagonal matrix

Note that it doesn't matter which sice of $\mathbf{D}$ is chosen as the inverse matrix, since they are inverses of one another no matter what way around they are placed.

Theorem 7.3. *An $n \times n$ matrix is diagonalizable iff its eigenequation has n linearly independent eigenvector solutions.*

Proving that diagonalizable matrixes have n linearly independent eigenvector solutions follows by reducing $\mathbf{P}^{-1}\mathbf{D}\mathbf{P}\mathbf{x} = \lambda\mathbf{x}$ to $\mathbf{D}(\mathbf{P}\mathbf{x}) = \lambda(\mathbf{P}\mathbf{x})$, then noting that this eigenequation has n linearly independent eigenvectors. To prove the other direction, we use our n eigenvectors and eigenvalues to construct such a diagonalizable matrix equal to $\mathbf{A}$.

Due to the constructive nature of this proof, we also obtain a method to calculate the diagonalized form of a diagonalizable matrix. This is known as *diagonalization.*

- $\mathbf{D}$ is the diagonal matrix of eigenvalues - $\mathbf{P}$ is the matrix of eigenvectors

Definition 7.4 (Multiplicity of an eigenvalue)**.** The *multiplicity of an eigenvalue* is the amount of linearly independent eigenvectors ar associated with said eigenvalue.

$$\det(\mathbf{P}^{-1}\mathbf{D}\mathbf{P}) = \prod_{i=1}^{n} \mathbf{D}_{i,i}$$

7.3.2 OrthMatrix diagonalization

Definition 7.5 (Orthogonally diagonalizable matrix)**.** An *orthogonally diagonalizable matrix* is a square matrix $\mathbf{A}$ such that there exists matrixes $\mathbf{P}, \mathbf{D}$ such that $\mathbf{A} = \mathbf{P}^\intercal \mathbf{D} \mathbf{P}$

- $\mathbf{P}$ is an orthogonal matrix
- $\mathbf{D}$ is a diagonal matrix

Theorem 7.4. A matrix is orhthogonally diagonalizable iff it is symmetric.

7.3.3 Properties of diagonalizable matrixes

7.3.4 Eigendecomposition

Chapter 8

Quadratic forms

$$q(\mathbf{x}) = \mathbf{x}^\intercal \mathbf{A} \mathbf{x}$$

$$q(\mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{A}_{ij} x_i x_j$$

Proposition 8.1. For any quadratic form q, there exists a unique symmetric matrix $\mathbf{A}$ such that q is representable as such.

$$q(\mathbf{x}) = \mathbf{x}^\intercal \mathbf{A} \mathbf{x}$$

Proposition 8.2. Let $q(\mathbf{x}) = \mathbf{x}^\intercal \mathbf{A} \mathbf{x}$ be a quadratic form, then $\frac{\mathbf{A}+\mathbf{A}^\intercal}{2}$ is a symmetric matrix that represents the same quadratic form

Theorem 8.1 (Principle axis theorem)**.** exists $\mathbf{yDy}$

Part II

Advanced

Chapter 9

Advanced matrix algebra

For a reader with a background in the fields of abstract algebra, further insight using group and ring theory can bring some concepts to light. This chapter is designed for those who are familiar with basic matrix theory and abstract algebra, and want to enrich their understanding of matrixes by combining the fields.

- $\mathrm{GL}(n, F)$ group - $\mathrm{SL}(n, F)$ group - $a\mathrm{O}(n, F)$ group - $\mathrm{SO}(n, F)$ group - Determinant as a group homomorphism from $\mathrm{GL}(n, F)$ to $F^\times$ - $\mathrm{GL}(n, F)/\mathrm{SL}(n, F)$ isomorphic to $F^\times$

9.1 Matrix exponential

It id defined in an analogous fashion to the exponential function of real and complex analysis.

Definition 9.1 (Matrix exponential)**.**

$$e^{\mathbf{X}} = \sum_{n=0}^{\infty} \frac{\mathbf{X}^n}{n!}$$

Proposition 9.1.

$$e^{\mathbf{X}} = \lim_{n\to\infty} (\mathbf{I} + \frac{\mathbf{X}}{n})^n$$

It arises in the study of simultaneous differential equations.

9.2 Cayley-Hamilton theorem

Definition 9.2 (Characteristic polynomial (scalar, linear algebra))**.**

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$$

Definition 9.3 (Characteristic polynomial (matrix, linear algebra))**.**

$$p_{\mathbf{A}}(\mathbf{B}) = \det(\mathbf{B} - \lambda\mathbf{I})$$

Chapter 10

Multilinear algebra

10.1 Multilinear maps

We have studied linear maps, and we have implicitly looked as some bilinear map; bivariate functions on two linear spaces that are linear (matrix multiplication and the inner product). Indeed there are many noteworthy multilinear maps aside from these, even if these two are the most renown.

Definition 10.1 (Multilinear map)**.** A *mulilinear map* or *multilinear transform* is a function $f : V^n \to W$

- f is additive in every argument

- f is homogeneous of degree 1 in every argument

This can be though of as a 'multivariate linear map'! Inner products, the cross product, and quadratic forms are examples of multilinear maps that we are already familiar with. We will now commence a deeper study on these operations generally.

10.1.1 Outer product

10.1.2 Hadamard product

10.1.3 Kronecker product

10.2 Tensor product

10.3 Grassmann algebra

Tensor theory will become fruitful to the study of the Grassman algebra

Definition 10.2 (Grassman algebra)**.** A *Grassman algebra over* $(V, K, +, \cdot)$ or *exterior algebra over* $(V, K, +, \cdot)$ is the quotient ring $\bigwedge(V) = T(V)/I$

- $T(V)$ is the tensor algebra on $(V, K, +, \cdot)$
- I is the ideal formed by $x \otimes x$ where $x \in V$

This definition relies heavily on ring theory, but what does it characterize? It is all about introducing the wedge product; an operation that generalizes the cross product to arbitrary linear spaces. The central rule encoded into the algebra is that $\mathbf{v} \times \mathbf{v} = \mathbf{0}$. It is not a commutative algebraic structure, however we can prove that it is anticommutative.

Wedge product k-blade bivector multivector Hodge star operator

Chapter 11

Tensors

The demands of physics have required more complicated classes of linear functions. The mathematical attempt to classify them has been by the invent of *tensors*.

We are at least familiar with scalars, vectors, linear maps, multilinear maps, linear forms, bilinear forms, and multilinear form. These are indeed all examples of different types of tensors.

11.1 Vectors and covectors (linear forms)

$[\mathbf{x}]_k$ means the scalar at row k of the column vector $\mathbf{x}$ $[\mathbf{A}]_{ij}$ means the scalar at row i, column j of the matrix $\mathbf{A}$

$T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ $\mathbf{f}_k = \mathbf{e}_i$

Consider a linear space V with an old basis $\mathbf{e}_i$ and new basis $\mathbf{f}_i$. We use some rule of the following form to translate each $\mathbf{e}_i$ to the new basis (allowing the transformation of any vector to the new basis).

$$\mathbf{f}_i = \sum_{j=1}^{n} [\mathbf{P}]_{ij} \mathbf{e}_j$$

$$\mathbf{e}_i = \sum_{j=1}^{n} [\mathbf{P}^{-1}]_{ij} \mathbf{f}_j$$

contravariant

$$\mathbf{P}^{-1}[\mathbf{v}]_{\mathcal{E}} = [\mathbf{v}]_{\mathcal{F}}$$

Vectors have the contravariant property
covariant

$$\mathbf{P}[\phi]_{\mathcal{E}} = [\phi]_{\mathcal{F}}$$

Linear forms (covectors) have the covariant property

11.2 Ricci calculus

Tensors are quite complex objects, so we introuce some notation that compactifys expressions while retaining all necessary information of the tensor.

Definition 11.1 (Tensor (indexed array definition))**.** A (p, q) *tensor* is a

11.2.1 Indexing laws

Superscript is used for contravariant coefficients Subscript is used for covariant coefficients

11.2.2 Einstein notation

Einsteun notation is essentially the omission of the Sigma summation symbol when the context is clear.

$$\sum_i \mathbf{e}_i x^i$$

The main idea is that *sums are implicitly implied over an index when you have a subscripted symbol times a superscripted symbol.* These expressions are so common that it becomes well worth introducing this notation.

the summation index may not appear more than twice to invoke Einstein notation.

11.3 Tensor

We have investigated scalars and vectors, and we've had the epiphany that matrixes represent linear maps on vectors; however just as we create an algebra upon matrixes (linear maps), can we form an algebra on 'tensors' (multilinear maps)?

Definition 11.2 (Tensor (multilinear map definition))**.** Let $(V, F, +, \cdot)$ be a linear space. A (p, q) *tensor* is a multilinear map that takes p covectors from V^* and q vectors from V as arguments, and outputs a scalar.

We know that a linear map can be defined pureply by knowing how it transforms a linear space's basis elements (hence why its matrix representation is possible).

The same holds for tensors, and it is why some think of tensors as a possibly multidimensional array (even if that explanation doesn't reveal a tensor's behaviour).

Proposition 11.1 (Tensors as an array)**.** Let $(V, F, +, \cdot)$ be a linear space with a (p, q) tensor T. Define the array of scalars $T^{i_1 \cdots i_p}_{j_1 \cdots j_q} = T(\varepsilon^{i_1}, \ldots, \varepsilon^{i_p}, \mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_q})$.

Definition 11.3 (Tensor algebra)**.**

Definition 11.4 (Tensor product)**.** For two linear spaces V, W, the *tensor product* is the linear space $V \otimes W$ associated with a bilinear map $T : V \times W \to V \otimes W$

We can create an equivalent but more abstract definition of a tensor in this way

Definition 11.5 (Tensor (tensor product definition))**.** A (p, q) tensor is an element of a tensor product space $\bigotimes_{n=1}^{p} V^* \otimes \bigotimes_{n=1}^{q} V$