

Math 247: Applied Linear Algebra ft. Dual Spaces

Shereen Elaidi

Winter 2018 Term

Contents

1	Introduction	2
2	Abstract Vector Spaces	3
2.1	Many Examples of Vector Spaces	4
2.2	Consequences of the Axioms	5
2.3	Subspaces	6
2.4	Linear Combination and Span	9
2.5	Linear Independence and Dependence	10
2.6	Applications (Theoretical) of Steinitz Exchange Lemma	11
3	Linear Maps	13
3.1	Endomorphisms	14
3.2	Coordinate Vectors	14
3.3	Representations of Linear Maps as Matrices	14
3.4	Composition of Linear Maps	14
3.5	Matrix Representations of Isomorphisms	15
3.6	Change of Basis	15
3.7	The Normal Form Problem	15
3.8	Characteristic Polynomial	16
3.9	Normal Form for Non-Diagonalisable Matrices	16
4	Applications of Linear Algebra (to other fields of math...not real life :-P)	17
4.1	Matrix Exponential	17
5	Orthogonality: A quick introduction	17
5.1	Standard inner product on \mathbb{R}^n	17
5.2	Properties of the Standard Inner Product on \mathbb{C}^n	18
5.3	Application of Orthogonality (Orthogonal and Unitary Matrices)	18
5.3.1	Orthogonal Case	18
5.3.2	Unitary Case	18
5.4	Orthogonality and Unitary: Diagonalisation and Upper Triangularisation	19
5.4.1	Spectral Theorem: Complex Case	19
5.5	Applications of the Real Spectral Theorem	20
5.5.1	Real Quadratic Forms	20
6	Inner Product Spaces	20
6.1	$K = \mathbb{R}$	21
6.2	Complex Inner Product Spaces	21
6.3	Geometry on a Vector Space	21
7	Fourier Series in 30 Minutes [not examinable]	22

1 Introduction

What is this class about?

1. Review of linear algebra, but with full proofs.
2. More abstract. We will move from \mathbb{R}^n to abstract vector spaces.
3. Many objects in maths behave similarly in the sense that they are actually instances of the same phenomenon. So, we want to generalise in order to...
 - a) Be more efficient.
 - b) Clarifies the essence of what's behind the object
 - c) Groups, rings, fields, and vector spaces are all such objects.

Definition 1 (Group). Let G be a set with a binary operation. A binary operation means that you can take two objects in the set, “add” them, and return another object in the set. A binary operation $G \times G \rightarrow G$ is a binary operation if it satisfies the following rules:

1. Law of Associativity: $\forall a, b, c \in G$:

$$a + (b + c) = (a + b) + c$$

$$\forall a, b, c \in G.$$

2. Existence of an additive neutral element: $\exists 0 \in G$ such that:

$$0 + a = a + 0 = a$$

$$\forall a \in G.$$

3. Existence of an additive inverse: $\forall a \in G$, there exists an element $-a \in G$ such that:

$$a + (-a) = (-a) + a = 0 \tag{1}$$

Any structure with a binary operation satisfying the above rules is called a group. A group always consists of two parts: a set and the operation. For example, a group equipped with addition is denoted $(G, +)$.

Sometimes, we would like to have a group whose binary operation is commutative. Not all groups are commutative!

Definition 2 (Abelian Group). A group $(G, +)$ is called abelian if $\forall a, b \in G$, we have the law of commutativity:

$$a + b = b + a$$

Example 1 (Examples of abelian groups). \mathbb{R} with standard addition, \mathbb{Z} with standard addition, \mathbb{Q} with standard addition, \mathbb{C} with standard addition. If we take n copies of \mathbb{R} and \mathbb{C} , denoted \mathbb{R}^n and \mathbb{C}^n , equipped with component-wise addition, then we also get groups.

However, \mathbb{N} with standard addition is not a group simply because there is *no* element with an additive inverse in \mathbb{N} for any value.

Example 2 (More examples of groups). 1. Matrix groups equipped with matrix addition. The sets of all $m \times n$ matrices with coefficients in \mathbb{R} and \mathbb{C} , respectively, are denoted as:

$$\text{MAT}(m \times n, \mathbb{R})$$

$$\text{MAT}(m \times n, \mathbb{C})$$

when equipped with component-wise addition they form a group.

2. Let's look at functions. The set $F(\mathbb{R})$ of all real-valued functions with domain \mathbb{R} , equipped with addition defined as:

$$(f + g)(x) := f(x) + g(x)$$

form a group. These have uncountably many components. The neutral element is the zero function:

$$0(x) := 0$$

and the additive inverse is:

$$(-f)(x) := -f(x) \quad \forall x \in \mathbb{R}$$

Example 3 (Non Abelian Group: General Linear). $\text{GL}(n, \mathbb{R})$ is called the general linear group. It is the set of all $n \times n$ invertible matrices with real coefficients together with the operation of matrix multiplication. The neutral element is the $n \times n$ identity matrix, denoted by I_n :

$$I_n A = A I_n = A \quad \forall A \in \text{GL}(n, \mathbb{R})$$

Since these matrices, the multiplicative inverse is A^{-1} , since $AA^{-1} = I_n$, which is the neutral element. Since associativity is clear, we have that $\text{GL}(n, \mathbb{R})$ forms a group.

Theorem 1 (Cancellation Law). Let $(G, +)$ be a group, and let $a, b, c \in G$ be elements such that:

1. $a + b = a + c \Rightarrow b = c$ and
2. $b + a = c + a \Rightarrow b = c$

Proof. We need to use the properties of a group as discussed. We will first prove (a). To that end, let's first use the existence of an additive inverse:

$$\begin{aligned} -a + (a + b) &= -a + (a + c) \\ \Rightarrow (-a + a) + b &= (-a + a) + c && \text{(Law of Associativity)} \\ \Rightarrow 0 + b &= 0 + c && \text{(By definition of neutral element)} \\ \Rightarrow b &= c && \text{(By neutral element law)} \end{aligned}$$

The proof for (b) is an exercise. □

2 Abstract Vector Spaces

Motivation: generalise \mathbb{R}^n .

Definition 3 (Vector Space over a field). Let V be a set, and let K be an arbitrary field (in this class we will assume that K is either \mathbb{R} or \mathbb{C}) together with two operations:

$$\begin{aligned} + : V \times V &\rightarrow V \quad \text{("Addition")} \\ \cdot : K \times V &\rightarrow V \quad \text{("Scalar multiplication")} \end{aligned}$$

and assume that $(V, +)$ is an abelian group and that the \cdot operation obeys the following rules:

1. "Associativity":

$$(k\ell)v = k(\ell v) \quad \forall k, \ell \in K, v \in V$$

2. "Distributivity of scalars over vectors":

$$(k + \ell)v = kv + \ell v \quad \forall k, \ell \in K, v \in V$$

3. “Distributivity of vectors over scalars”:

$$k(u + v) = ku + kv \quad \forall k \in K, u, v \in V$$

4. “The existence of the neutral element of scalar multiplication”

$$1 \cdot v = v \quad \forall v \in V$$

Then, the tuple $(V, +, \cdot)$ is called a vector space over K .

Remark: the final axiom is not self-evident. In fact, it is not provable from other axioms. For more details see Assignment 1.

2.1 Many Examples of Vector Spaces

Example 4 (Examples of vector spaces). \mathbb{R}^n endowed with standard vector addition and scalar multiplication, denoted $(\mathbb{R}^n, +, \cdot)$ forms a vector space. Additionally, the set of all $n \times m$ matrices with component-wise addition and standard scalar multiplication, denoted by $(\text{MAT}(m \times n, \mathbb{K}), +, \cdot)$ also forms a vector space over K .

Example 5 (Sequence Vector Spaces). Let S be the set of all real-valued sequences, i.e.:

$$\{(a_1, a_2, a_3, \dots) \mid a_i \in \mathbb{R}\}$$

This can be considered a vector space. Addition is component-wise and multiplication is also component-wise. It is similar to \mathbb{R}^n , except for \mathbb{R}^n we stop at some finite number n . Thus, this space of sequences is *infinite-dimensional*. It is trivially a vector space over \mathbb{R} , since checking the axioms amounts to passing them down to the component-level, since all operations are component-wise, and those will trivially follow from the fact that the components are elements of the vector space $(\mathbb{R}, +, \cdot)$.

The following examples provide interesting vector spaces for the field of mathematics real analysis:

Example 6 (Spaces of convergent real-valued sequences). Let c be the set of all convergent real sequences together with component-wise addition and scalar multiplication. We need to be careful, since something can go wrong. In particular, the $+$ operation needs to be a binary operation, meaning that when you add two members of the set you obtain another member in the set. We will use and not prove the following result from real analysis: the sum of two convergent sequences converge and they converge to the limits of the individual sequences that we add. Thus, the only thing we need to check is that c is closed under \cdot and $+$. However, both of these results are proven in analysis.

Example 7 (The space of all real-valued sequences that converge to zero c_0). This space is denoted by c_0 and is clearly closed under $+$ and \cdot . This forms a subspace of c and is also infinite-dimensional but it is not countable.

Example 8 (The space c_{00} , the space of all real sequences that are eventually zero). c_{00} . This space contains sequences where all but finitely many terms are different from zero. This is clearly closed under addition and multiplication, and it is a subspace of c_0 . While this vector space is still infinite dimensional, it is countable.

Example 9. Let $I \subseteq \mathbb{R}$ be an interval. Consider $F(I)$. This is the set of all real-valued functions on I with $+$ and \cdot defined by:

$$\begin{aligned}(f + g)(x) &= f(x) + g(x) \\ (kf)(x) &= kf(x)\end{aligned}$$

This example will be investigated in Assignment 1. It is a vector space.

There are a lot of functions that mathematicians care about: differentiable functions, continuous functions, those that have Taylor expansions on I , etc. These are all vector spaces.

Example 10. (a) Let $F^c(I)$ be the set of all continuous functions on I . (b) Let $C^0(I)$ be the set of all continuously differentiable functions on I . (c) Let $C^n(I)$ be the set of all functions on I with continuous derivatives up to order n . (d) Let $C^\infty(I)$ be all functions on I with derivatives of all order. Define operations on (a) - (d) as was defined in the previous example. These are all vector spaces over \mathbb{R} .

Linear algebra is also useful for solving differential equations. For example, consider the set of all solutions of $y'' + y = 0$, or more generally, the set of all solution to any linear differential equation:

$$a_n y^{(n)} + a_{(n-1)} y^{(n-1)} + \dots + a_n y' + a_0 y = 0$$

where $a_0, \dots, a_n \in \mathbb{R}$. This is an important set of functions. We really care about the closedness of the so-called solution space. So, we need to verify that solutions are closed under $+$ and \cdot . To that end, let y_1, y_2 be solutions of $y'' + y = 0$. Then we have:

$$y_1'' + y_1 = 0 \text{ and } y_2'' + y_2 = 0$$

Now let's study $(y_1 + y_2)'' + (y_1 + y_2)$. By the linearity of differentiation:

$$y_1'' + y_2'' + y_1 + y_2 = (y_1'' + y_1) + (y_2'' + y_2) = 0 + 0 = 0$$

Hence, we have closure under addition. For scalars:

$$(ky_1)'' + ky_1 = ky_1'' + ky_1 + k(y_1'' + y_1) = 0$$

which implies that ky_1 is also a solution, and we therefore have closure under scalar multiplication.

Once we determine that the solutions are a vector space, then all we need to do is obtain a basis and take all *linear combinations* to obtain more solutions of DEs.

2.2 Consequences of the Axioms

This section has some immediate consequences from the axioms of a vector space.

Theorem 2. Let $(V, +, \cdot)$ be a vector space over K . Then:

1. $k \cdot 0 = 0 \forall k \in K$.
2. $0 \cdot v = 0 \forall v \in V$.
3. $k \cdot v = 0 \iff k = 0 \text{ or } v = 0$.

Proof. 1. We need to show that $k \cdot 0 = 0$. This is a "fun trick proof."

$$\begin{aligned} k \cdot 0 &= k \cdot (0 + 0) && (0 \text{ is the neutral element of } +) \\ k \cdot 0 &= k \cdot 0 + k \cdot 0 && (\text{distributivity of scalars over vectors}) \\ 0 + k \cdot 0 &= k \cdot 0 + k \cdot 0 && (\text{by the neutral element law}) \\ 0 &= k \cdot 0 && (\text{by the cancellation law}) \end{aligned}$$

2.

$$\begin{aligned} 0 \cdot v &= (0 + 0)v && (0 \text{ is neutral element of addition in } K) \\ 0 \cdot v &= 0v + 0v && (\text{distributivity of scalars over vectors}) \\ 0 + 0 \cdot v &= 0 \cdot v + 0 \cdot v && (\text{neutral element law}) \\ 0 &= 0 \cdot v && (\text{cancellation law}) \end{aligned}$$

3. There are two directions to show here.

" \Leftarrow ": if $k = 0$ or $v = 0$, then it follows from (a) + (b), respectively, that $kv = 0$.

“ \Rightarrow ”: there are two cases to consider. If $k = 0$, then we are done. If $k \neq 0$, then we need to prove that $v = 0$. Take $kv = 0$. It has a multiplicative inverse:

$$kv = 0 \Rightarrow \left(\frac{1}{k}\right)(kv) = \frac{1}{k}0 = 0$$

Apply associativity to the $\left(\frac{1}{k}\right)(kv)$ term:

$$\left(\frac{1}{k}\right)v \underbrace{=}_{\text{axiom}} 1v = v$$

This implies that $v = 0$, and so we are done. □

Theorem 3. Let $(V, +, \cdot)$ be a vector space over K . Then:

1. $n \cdot v = \underbrace{v + \dots + v}_{n \text{ times}}$ for all $n \in \mathbb{N}$.
2. $0 \cdot v = 0$.
3. $(-1)v = -v$ (the additive inverse) for all $v \in V$.

Proof. 1. We will prove this one by induction. The base case is obvious since it is an axiom of a vector space ($1v = v$). Now for the inductive step assume that

$$nv = \underbrace{v + \dots + v}_{n \text{ times}}$$

for some n in \mathbb{N} . Applying distributivity:

$$(n+1)v = nv + 1v = \underbrace{v + \dots + v}_{n \text{ times, by IH}} + \underbrace{v}_{\text{by axiom}} = \underbrace{v + \dots + v}_{n+1 \text{ times}}$$

where the final equality follows from the fact that we are in an abelian group.

2. We get this from (1); we have v , zero times.
3. We want to prove that $(-1)v$ is the additive inverse. We can simply add this candidate and prove that the sum is zero:

$$\begin{aligned} v + (-1)v & \qquad \qquad \qquad \text{(use the axiom)} \\ (1)v + (-1)v & \qquad \qquad \qquad \text{(use the other axiom)} \\ (1 + (-1))v & \qquad \qquad \qquad \text{(distributivity)} \\ 0v & \end{aligned}$$

and we know that $0v = 0$ by (2). We thus have that $(-1)v$ is the additive inverse of v , which is what we wanted to show. □

Exercise: prove that $(-n)v = \underbrace{(-v) + \dots + (-v)}_{n \text{ times}}$.

2.3 Subspaces

Definition 4 (Subspaces). Let $(V, +, \cdot)$ be a vector space over K . Let $W \subseteq V$. Then, we say that W is a subspace of V , in symbols $W \leq V$, if $(W, +, \cdot)$ is a vector space over K in its own right.

Q: How can we determine if something is a subspace? This question motivates the next theorem.

Theorem 4. Let $(V, +, \cdot)$ be a vector space over K , $W \subseteq V$. Then, $W \leq V$ if and only if:

1. W is closed under addition and multiplication.

2. $0 \in W$.

Note that the final requirement in the theorem is actually a proxy for requiring that W is not the empty set.

Proof. “ \Rightarrow ”: Since W is a vector space over K , W is closed under $+$ and \cdot . Furthermore, $0 \in W$ since $(W, +)$ is an abelian group and thus contains the neutral element.

“ \Leftarrow ”: This is a straightforward verification of the axioms of a vector space. Most of the axioms will simply follow, since as a subset of V W will inherit a lot of rules.

- W will inherit associativity from V .
- W will inherit commutativity from V .
- W contains the neutral element by condition.
- Let u be an arbitrary vector. Since U is closed under scalar multiplication, we have that $(-1)u = -u \in U$. This proves the existence of an additive inverse.
- All of the scalar multiplication axioms will hold since they hold $\forall v \in V$.

This implies that W is a vector space with respect to $+$ and \cdot , which gives us that $W \leq V$, which is what we wanted to show. \square

Here are some examples of subspaces of vector spaces.

Example 11. The following are the *only* subspaces of \mathbb{R}^2 :

1. $\{0\}$
2. All lines through the origin.
3. The whole of \mathbb{R}^2 .

Example 12. Let $V := \text{MAT}(m \times n, \mathbb{K})$. Let $U :=$ the set of all upper triangular matrices i.e.:

$$U := \{ A = (a_{ij}) \in \text{MAT}(m \times n, \mathbb{K}) \mid a_{ij} = 0 \forall i > j, 1 \leq i, j \leq n \}$$

This is a subspace.

Example 13. Recall sequence spaces. We have the following chain of subspaces: $C_{00} \leq C_0 \leq C \leq S$, where S denotes a sequence space.

Example 14. Let $P_n(K)$ be the set of polynomials of degree at most n . Let $P(K)$ be the set of all polynomials, and equip it with standard addition of polynomials. Then:

$$P_0(K) \leq P_1(K) \leq P_2(K) \leq \dots \leq P_n(K) \leq \dots \leq P$$

The dimension of each P_n is $n + 1$.

Theorem 5. Let $(V, +, \cdot)$ be a vector space over K .

1. Let U be a subspace of V and let W be a vector space. Then: $U \cap W \leq V$.
2. Let $U_1, \dots, U_n \leq V$. Then: $U_1 \cap \dots \cap U_n = \bigcap_{i=1}^n U_i \leq V$.
3. Let I be an arbitrary index set, and let $U_i \leq V \forall i \in I$. Then: $\bigcap_{i \in I} U_i \leq V$.

Proof. 1. Another straightforward verification of the axioms.

a) $0 \in U$, and $0 \in W$ since W is a vector space. Thus $0 \in W \cap U$.

b) Now we check closure under addition and scalar multiplication.

- i. Let $v_1, v_2 \in U \cap W$. We thus especially have that $v_1 \in U$ and $v_2 \in U$, and so $v_1 + v_2 \in U$ since U is closed under addition. Similarly, $v_1 + v_2 \in W$. Since the sum is in both sets, it is in the intersection of the sets. Hence, it is closed under addition.

- ii. Since $u_1 \in U$, $Ku_1 \in V$ for all $k \in K$, since U is a vector space. Similarly, if $u_2 \in W$, then $ku_2 \in W$ for all $k \in K$. Thus, the product is in the intersection.

So, by the subspace criteria, $U \cap W \leq V$.

2. Exercise.
3. Exercise.

□

So, intersections are generally not a problem. However, unions are a more difficult case. For example let $V = \mathbb{R}^2$, and define U to be the x -axis and W to be the y -axis. Consider $U \cup W$. It is not a line or plane, so it is not a subspace. While this union of subspaces will always contain zero and while it will always be closed under scalar multiplications, it is not closed under addition. Unions in general are not subspaces, which motivates the following theorem. The proof is in assignment two.

Theorem 6. Let $(V, +, \cdot)$ be a vector space over K . Then, $U \cup W \leq V$ if and only if $U \subseteq W$ or $W \subseteq U$ (i.e., the trivial cases).

Proof. Homework. □

Definition 5 (Sum of Vector Spaces). Let $(V, +, \cdot)$ be a vector space over K . Let $U, W \leq V$. Then:

$$U + W := \{u + w \mid u \in U, w \in W\}$$

which is called the sum of U and W , is defined as follows: all the possible sums with one vector from U and one vector from W .

Q: Why do we bother defining the sum?

1. It is a subspace of V .
2. It is the *smallest* subspace of V that contains the union of U and W (“generator”).

Theorem 7. Let $(V, +, \cdot)$ be a vector space over K . Let $U, W \leq V$. Then:

1. $U + W \leq V$.
2. $U + W$ is the *smallest* subspace of V that contains both U and W and thus their union. This means that if \tilde{V} is any subspace of V with $U \cup W \subseteq \tilde{V}$, then the sum $U + W \subseteq \tilde{V}$.

Proof. 1. Another straightforward verification of the axioms.

- a) We have that $0 \in U$ and $0 \in W$, and so $0 = 0 + 0 \in U + W$.
- b) Let $v_1, v_2 \in U + W$. Then, $\exists u_1, u_2 \in U, w_1, w_2 \in W$, such that:

$$\begin{aligned} v_1 &= u_1 + w_1 \\ v_2 &= u_2 + w_2 \end{aligned}$$

Adding these together we obtain:

$$v_1 + v_2 = \underbrace{(u_1 + u_2)}_{\in U} + \underbrace{(w_1 + w_2)}_{\in W} \in U + W$$

- c) Now for scalar multiplication. Write

$$\begin{aligned} kv_1 &= k(u_1 + w_1) \\ &= ku_1 + kw_1 \end{aligned} \quad (\text{distributivity})$$

$ku_1 \in U$ and $kw_1 \in W$, which means that $kv_1 \in U + W$, which is what we wanted to show. Therefore, $U + W \leq V$.

2. Let \tilde{V} be any subspace of V that contains both U and W . Let $u \in U$ and $w \in W$ be arbitrary. Then, by definition, $u + w \in U + W$. Every element must thus be in \tilde{V} , which implies that $U + W$ must be contained in \tilde{V} , which proves that $U + W$ is the smallest subspace of V containing $U \cup W$. □

In a similar fashion we can define $U_1 + \dots + U_n$, where U_1, \dots, U_n are all subspaces of V as:

$$\{u_1 + \dots + u_n \mid u_k \in U_k \forall 1 \leq k \leq n\}$$

Then it's easy to verify that $U_1 + \dots + U_n \leq V$.

2.4 Linear Combination and Span

Definition 6 (Linear Combination). Let $(V, +, \cdot)$ be a vector space over K . Let $v_1, \dots, v_n \in V$ and let $k_1, \dots, k_n \in K$. Then, an expression of the form

$$k_1v_1 + \dots + k_nv_n \quad (2)$$

is called a **linear combination** of v_1, \dots, v_n .

Note that it is possible to extend this to infinitely many vectors:

$$k_1v_1 + k_2v_2 + \dots$$

but this is *completely meaningless*. This is a bad thing since we do not know what it means to add infinitely many vectors. We have not developed a notion of convergence or distance on a vector space in this class, so we will not discuss this. What we can do is write $k_1v_1 + k_2v_2 + \dots$ where all but *finitely many* of the $k_i = 0$, but this can be viewed as an abuse of notation, since the sum is actually finite.

Definition 7 (Span). Let $(V, +, \cdot)$ be a vector space over K . Let $v_1, \dots, v_n \in V$. Then:

$$\text{span}\{v_1, \dots, v_n\} := \{k_1v_1 + \dots + k_nv_n \mid 1 \leq i \leq n, k_i \in K\} \quad (3)$$

is called the **span** of v_1, \dots, v_n . It is the set of all possible linear combinations. We can extend this to infinitely many vectors $v_1, v_2, \dots \in V$ as:

$$\text{span}\{v_1, \dots, v_n\} := \{k_1v_1 + k_2v_2 + \dots \mid k_i \in K, \text{ all but finitely many } k_i \text{ are zero}\}$$

Note that

$$\text{span}\{v_1, \dots, v_n\} = \{v_1\} + \dots + \{v_n\}$$

Which means that the span of v_1, \dots, v_n can be written as:

$$\text{span}\{v_1, \dots, v_n\} = \{k_1v_1 + \dots + k_nv_n \mid k_i \in K\}$$

From this, we also obtain that the span of the vectors v_1, \dots, v_n is the *smallest* subspace of V containing the vectors v_1, \dots, v_n .

Example 15. Let $V = \mathbb{R}^n$. Then, the standard basis vectors: $e_1 = (1, 0, 0, \dots)$, $e_2 = (0, 1, 0, \dots)$, \dots , $e_n = (0, 0, \dots, 1)$ spans \mathbb{R}^n .

Example 16. Consider the vector space $P_n(K)$. The building blocks of this vector space are:

$$P_n(K) = \{1, x, x^2, x^3, \dots, x^n\}$$

We thus have:

$$\begin{aligned} \text{span}\{1, x, x^2, \dots, x^n\} &= \{k_0 + k_1x + \dots + k_nx^n \mid k_i \in K, 0 \leq i \leq n\} \\ &= P_n(K) \end{aligned}$$

Example 17. Consider the vector space S of all sequences with coefficients in \mathbb{R} . Define:

$$\begin{aligned} s_1 &= (1, 0, 0, \dots) \\ s_2 &= (0, 1, 0, \dots) \\ s_3 &= (0, 0, 1, 0, \dots) \\ &\vdots \end{aligned}$$

This is similar to the previous example where we had the standard basis of \mathbb{R}^n . However, $\{s_1, s_2, \dots\}$ does NOT span s since linear combinations only involve finitely many terms. Every linear combination of s results in a sequence that is eventually *constantly equal to zero*, i.e., the subspace c_{00} .

2.5 Linear Independence and Dependence

Definition 8 (Linear Independence and Dependence). Let $(V, +, \cdot)$ be a vector space over K , $v_1, \dots, v_n \in V$. We say that v_1, \dots, v_n are linearly independent if

$$k_1 v_1 + \dots + k_n v_n = 0$$

implies that $k_1 = k_2 = \dots = k_n = 0$. v_1, \dots, v_n are called linearly dependent if they're not linearly independent.

Theorem 8. Let $(V, +, \cdot)$ be a vector space over K . Then:

1. Any (finite) subset of V that contains the zero vector is linearly dependent.
2. If v_1, \dots, v_n are linearly dependent, then at least one of them, say v_j , can be written as a linear combination of the remaining vectors. In that case, consider $\text{span}\{v_1, \dots, v_n\}$. Then we can kick v_j out and still have the same span:

$$\text{span}\{v_1, \dots, v_n\} = \text{span}\{v_1, \dots, \hat{v}_j, \dots, v_n\}$$

Proof. 1. Consider $0, v_1, \dots, v_n$. This is a linearly dependent set and we can prove it by providing a non-trivial combination that produces zero:

$$1 \cdot 0 + 0v_1 + \dots + 0v_n = 0$$

since the coefficient of the zero vector is non-zero. **Remark:** even just the set with the zero vector, $\{0\}$, is linearly dependent. This is because $1 \cdot 0 = 0$, which is a non-trivial combination of one vector resulting in the zero vector.

2. Let v_1, \dots, v_n be linearly dependent. By definition, there exists scalars $k_1, \dots, k_n \in K$, not all of them zero, such that:

$$k_1 v_1 + \dots + k_j v_j + \dots + k_n v_n = 0$$

Assume that $k_i \neq 0$. Then, we can solve the equation:

$$k_j v_j = -k_1 v_1 - \dots - k_{j-1} v_{j-1} - \dots - k_{j+1} v_{j+1} - \dots - k_n v_n$$

Since $k_j \neq 0$, we can divide by k_j :

$$v_j = \frac{-k_1}{k_j} v_1 - \dots - \frac{-k_n}{k_j} v_n$$

Which is a linear combination of the remaining vectors that results in v_j . We now want to show that the spans with and without v_j are the same. Obviously, $\text{span}\{v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n\}$ is a subset of $\text{span}\{v_1, \dots, v_n\}$. We need to prove the opposite inclusion:

$$\text{span}\{v_1, \dots, v_n\} \subseteq \text{span}\{v_1, \dots, \hat{v}_j, \dots, v_n\}$$

To that end, take an arbitrary element of the space of v_1, \dots, v_n :

$$v = a_1 v_1 + \dots + a_j v_j + \dots + a_n v_n \in \text{span}\{v_1, \dots, v_n\}$$

Then, we can write this without v_j by what we did in the above when we solved for v_j ;

$$v = a_1 v_1 + \dots + a_j \left(\frac{-k_1}{k_j} v_1 - \dots - \frac{-k_n}{k_j} v_n \right) + \dots + a_n v_n$$

Collecting the coefficients:

$$v = \left(a_1 - \frac{k_1}{k_j} a_j \right) v_1 + \dots + \left(a_{j-1} - \frac{k_{j-1}}{k_j} a_j \right) v_{j-1} + \left(a_{j+1} - \frac{k_{j+1}}{k_j} a_j \right) v_{j+1} + \dots + \left(a_n - \frac{k_n}{k_j} a_j \right) v_n$$

Which is a linear combination of $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n$. □

Theorem 9. Let $v_1, \dots, v_n \in V$, v_1, \dots, v_n be linearly independent. Let $v_{n+1} \notin \text{span}\{v_1, \dots, v_n\}$. Then, v_1, \dots, v_n, v_{n+1} are linearly independent.

Proof. To show this, we will take an arbitrary linear combination that results in zero, and show that this means that the coefficients must be zero. To that end, let

$$k_1 v_1 + \dots + k_n v_n + k_{n+1} v_{n+1} = 0$$

be an arbitrary linear combination. We will use the linear independence of v_1, \dots, v_n to show that $k_1, \dots, k_n = 0$. For k_{n+1} , for a contradiction, assume that $k_{n+1} \neq 0$. Then, solve for v_{n+1} :

$$\begin{aligned} k_{n+1} v_{n+1} &= -k_1 v_1 - \dots - k_n v_n \\ \Rightarrow v_{n+1} &= -\frac{k_1}{k_{n+1}} v_1 - \frac{k_n}{k_{n+1}} v_n \end{aligned}$$

Which is impossible, since we assumed that $v_{n+1} \notin \text{span}\{v_1, \dots, v_n\}$. This proves that $k_1, \dots, k_n, k_{n+1} = 0$, which shows that v_1, \dots, v_n must be linearly independent. \square

Definition 9 (Finite Dimensional). Let V be a vector space over K . V is called **finite-dimensional** if V has a finite spanning set i.e., there are finitely many $v_1, \dots, v_n \in V$ such that they all span V :

$$V = \text{span}\{v_1, \dots, v_n\}$$

Remark: Let $(V, +, \cdot)$ be a vector space over K . Assume $W \leq V$. Assume that V is finite-dimensional. It is tempting to assume that W is finite-dimensional; this is true, but we do not have any tools to prove it at the moment. This will be proven later.

Definition 10 (Infinite-Dimensional). V is called **infinite-dimensional** if it is not finite dimensional.

Methods of showing that a vector space is infinite dimensional are very different from methods of showing that a vector space is finite-dimensional; for finite-dimensional, all we need to do is construct one example of a finite spanning set. However, to prove a vector space is infinite dimensional requires showing that no finite spanning set can possibly exist.

Example 18. \mathbb{R}^n is finite-dimensional. This is trivial; write out the following finite spanning set:

$$\mathbb{R}^n = \text{span}\{e_1, \dots, e_n\}$$

Example 19. c_{00} is infinite-dimensional. We will do a proof by contradiction. Assume that c_{00} is finite-dimensional. Let s_1, \dots, s_n be sequences in c_{00} such that $c_{00} = \text{span}\{s_1, \dots, s_n\}$.

No more proofs; just summary.

Theorem 10 (Steinitz Exchange Lemma). Let V be a finite-dimensional vector space over K . Let $u_1, \dots, u_m \in V$ be linearly independent. Let v_1, \dots, v_n be spanning (i.e., $V = \text{span}\{v_1, \dots, v_n\}$). Then, $m \leq n$ and there exist finitely many indices k_1, \dots, k_{n-m} such that

$$u_1, \dots, u_m, v_{k_1}, \dots, v_{k_{n-m}} \tag{4}$$

is spanning (i.e., we exchange m of the vectors v_1, \dots, v_n by u_1, \dots, u_m without changing the span.)

2.6 Applications (Theoretical) of Steinitz Exchange Lemma

Theorem 11. Every subspace of a finite-dimensional vector space is finite-dimensional.

Definition 11 (Basis). Let V be a vector space over K . A subset \mathcal{B} of V is called a **basis** of V if:

1. \mathcal{B} is linearly independent.
2. \mathcal{B} is spanning.

Example 20. For \mathbb{R}^n , $\mathcal{B} = \{e_1, \dots, e_n\}$.

Example 21. For P_n (recall these are polynomials of degree less than or equal to n), the basis would be $\mathcal{B} = \{1, x, \dots, x^n\}$.

Example 22. Set $V = \text{MAT}(m \times n, \mathbb{K})$. Let

$$M_{ij} : (m_{ij}) \text{ where } m_{kl} \text{ is } \begin{cases} 0 & \text{if } kl \neq ij \\ 1 & \text{if } kl = ij \end{cases} \quad (5)$$

Then, M_{ij} is a basis for V .

Theorem 12. Every finite-dimensional vector space has a basis.

Theorem 13. Let V be a finite-dimensional vector space over K . Let $\{u_1, \dots, u_m\}$ and $\{v_1, \dots, v_n\}$ be two bases for V . Then $m = n$, i.e., any two bases of a finite-dimensional vector space contain the same number of elements.

Definition 12 (Dimension). Let V be a finite-dimensional vector space over K . Let $\{v_1, \dots, v_n\}$ be any basis for V . Then, n is the **dimension** of V .

Example 23.

1. \mathbb{R}^n as a vector space over \mathbb{R} has dimension n .
2. \mathbb{C}^n as a vector space over \mathbb{C} has dimension n .
3. $P_n(K)$ as a vector space over K has dimension $k + 1$.
4. $\text{MAT}(m \times n, \mathbb{K})$ has dimension $n \times m$.

Theorem 14. Let V be a finite-dimensional vector space over K . Let $V = \text{span}\{B_1, \dots, B_n\}$ be a basis for V . Let $v \in V$. Then, there exists a uniquely determined $k_1, \dots, k_n \in K$ such that

$$v = k_1 v_1 + \dots + k_n v_n \quad (6)$$

Theorem 15. Let V be a finite-dimensional vector space over K . Let $U \leq V$. then, $\dim(U) \leq \dim(V)$.

Theorem 16. This theorem answers the question: when can we stop constructing bases? Let V be a finite-dimensional vector space over K of dimension n . Let $\{v_1, \dots, v_n\}$ be a set of n linearly independent vectors. Then, $\{v_1, \dots, v_n\}$ is a basis for V .

Theorem 17. Let V be a finite-dimensional vector space over K of dimension n . Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be spanning. Then, \mathcal{B} is a basis for V .

Theorem 18. Let V be a finite-dimensional vector space over K . Let $U \leq V$ and let $\mathcal{B}_u := \{u_1, \dots, u_m\}$ be a basis for U . Then, \mathcal{B}_u can be extended to a basis for V i.e., $\exists u_{m+1}, \dots, u_n \in V$ such that

$$\mathcal{B}_v := \{u_1, \dots, u_m, u_{m+1}, \dots, u_n\} \quad (7)$$

is a basis for V .

Theorem 19. Let V be a finite-dimensional vector space over K . Let $U \leq V$ and $W \leq V$. Then:

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W) \quad (8)$$

Example 24. Any two planes through the origin in \mathbb{R}^3 have non-trivial intersection, i.e., they intersect in at least a line through the origin. One can use the above theorem to give a formal proof of this.

Definition 13 (Direct Sum). Let $(V, +, \cdot)$ be a vector space over K . Let $U \leq V, W \leq V$. Define:

$$U + W := \{u + w \mid u \in U, w \in W\} \quad (9)$$

Now, the sum is called **direct** if the following property holds: $\forall v \in U + W$, the representation $v = u + w$, where $u \in U$ and $w \in W$ is uniquely determined. We then write $V = U \oplus W$.

Theorem 20. A sum $U + W$ of subspaces of V is direct $\iff U \cap W = \{0\}$.

3 Linear Maps

In mathematics one often wants to study maps between objects that preserve structures. In our case, we want to study linear maps.

Definition 14 (Linear). Let V and W be vector spaces over K (it must be the same K). A function $L : V \rightarrow W$ is called linear if it behaves well with respect to the operations that we have:

1. $L(v_1 + v_2) = L(v_1) + L(v_2)$.
2. $L(kv_1) = kL(v_1) \forall k \in K$.

Theorem 21. Let $L : V \rightarrow W$ be linear. Then the zero vector is preserved:

$$L(0) = 0 \tag{10}$$

Caution! in calculus the so called “linear” maps $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto ax + b$ is only linear if $b = 0$. This is similar to the fact that lines in general are not subspaces.

Definition 15 (Subspaces Associated to Linear Maps: Kernel and Image). Let $L : V \rightarrow W$ be linear. Then:

1. The kernel is defined as:

$$\ker(L) := \{v \in V \mid L(v) = 0\} \subseteq V \tag{11}$$

2. The image is defined as:

$$\text{im}(L) := \{L(v) \mid v \in V\} \subseteq W \tag{12}$$

Theorem 22. Let $L : V \rightarrow W$ be linear. Then:

1. The kernel is a subspace of V .
2. The image is a subspace of W .

Example 25. Let $A \in \text{MAT}(m \times n, \mathbb{K})$. Then, $\ker[A]$ is the null space of A and $\text{im}[A]$ is the column space of A .

Example 26. Consider the derivative map: $\mathcal{D} : P(\mathbb{R}) \rightarrow P(\mathbb{R}), p \mapsto p'$. Then $\ker[\mathcal{D}] = P_0(\mathbb{R})$ (the set of all constant polynomials) and $\text{im}[\mathcal{D}] = P(\mathbb{R})$ (everything).

Theorem 23. Let V, W be vector spaces over K . Let $L : V \rightarrow W$ be linear. Then L is injective $\iff \ker[L] = \{0\}$.

Theorem 24. This theorem tells us how we can construct linear maps. Let V be a finite-dimensional vector space over K , let W be any vector space over K . Let $\{v_1, \dots, v_n\}$ be a basis for V . Let $\{w_1, \dots, w_n\}$ be arbitrary vectors in W . Then, there exists a uniquely determined linear map $L : V \rightarrow W$ such that

$$L(v_1) = w_1, \dots, L(v_n) = w_n \tag{13}$$

This tells us that we can uniquely determined a linear map by its basis.

Theorem 25 (Dimension Formula – Fundamental Theorem). Let V be a finite-dimensional vector space over K , W a vector space over K , $L : V \rightarrow W$ a linear map.

$$\dim \ker[L] + \dim \text{im}[L] = \dim V \tag{14}$$

3.1 Endomorphisms

Definition 16 (Endomorphism). A linear map $L : V \rightarrow V$ is called a endomorphism.

Theorem 26. Let V be a finite-dimensional vector space over K , $L : V \rightarrow V$ an endomorphism. Then

$$L \text{ bijective} \iff L \text{ injective} \iff L \text{ surjective} \quad (15)$$

Lemma 27. Let V be a finite-dimensional vector space over K and let $U \leq V$ such that $\dim U \leq \dim V$. Then $U = V$.

3.2 Coordinate Vectors

Theorem 28. Let V be a finite-dimensional vector space and let \mathcal{B} be an ordered basis for V . Then $[\cdot]_{\mathcal{B}} : V \rightarrow K^n$, $v \mapsto [v]_{\mathcal{B}}$ is linear and bijective, i.e. such a map is called a isomorphism.

Theorem 29. Let $L : V \rightarrow W$ be an isomorphism. Then, L has an inverse map $L^{-1} : W \rightarrow V$ which is linear.

Theorem 30. Let V be a vector space over K of dimension n . Then, V is isomorphic to K^n .

3.3 Representations of Linear Maps as Matrices

Theorem 31. Let $L : V \rightarrow W$ be linear, $\dim V = m$, $\dim W = n$. Let $\mathcal{B} = \{v_1, \dots, v_m\}$ be a basis for V and let $\mathcal{B}' = \{w_1, \dots, w_n\}$ be a basis for W . Then

$$[L(v)]_{\mathcal{B}'} = [L]_{\mathcal{B}', \mathcal{B}} [v]_{\mathcal{B}} \quad (16)$$

$\forall v \in V$.

Theorem 32. Let $L : V \rightarrow W$ be linear, where V and W are finite-dimensional. Let \mathcal{B} be a basis for V and let \mathcal{B}' be a basis for W and define $A := [L]_{\mathcal{B}', \mathcal{B}}$. Then:

1. $\ker[L] \simeq \text{nullspace}(A)$.
2. $\text{im}[L] \simeq \text{colspace}(A)$.

the isomorphism is given by the following map:

$$\ker[L] \rightarrow \text{nullspace}(A) \quad (17)$$

$$v \mapsto [v]_{\mathcal{B}} \quad (18)$$

and

$$\text{im}[L] \rightarrow \text{colspace}(A) \quad (19)$$

$$w \mapsto [w]_{\mathcal{B}'} \quad (20)$$

3.4 Composition of Linear Maps

Theorem 33. Let $L_1 : U_{\mathcal{B}} \rightarrow V_{\mathcal{B}'}$ and $L_2 : V_{\mathcal{B}''} \rightarrow W_{\mathcal{B}''}$ be two linear maps. Let $\dim U = m$, $\dim V = k$, and $\dim W = n$. Then:

$$\underbrace{[L_2 \circ L_1]_{\mathcal{B}'', \mathcal{B}}}_{n \times m} = \underbrace{[L_2]_{\mathcal{B}'', \mathcal{B}'}}_{n \times k} \circ \underbrace{[L_1]_{\mathcal{B}', \mathcal{B}}}_{k \times m} \quad (21)$$

(Remark: this is why matrix multiplication in Math 133 was defined in such a complicated way).

3.5 Matrix Representations of Isomorphisms

Theorem 34. Let U, V be finite-dimensional vector spaces over a field K . Let $L : V \rightarrow W$ be linear and let $\mathcal{B}, \mathcal{B}'$ be bases for V, W , respectively. Then, L is an isomorphism \iff the matrix representation map $[L]_{\mathcal{B}', \mathcal{B}}$ is invertible.

3.6 Change of Basis

Two driving questions of this section:

1. If V is a finite-dimensional vector space over K , $\mathcal{B}, \mathcal{B}'$ are bases for V , then how are the coordinate vectors of $v \in V$ related for each of the bases? In other words, let $v \in V$. Given $[v]_{\mathcal{B}}$, how can one compute $[v]_{\mathcal{B}'}$? We call this a **change of basis**.
2. Let $L : V \rightarrow V$ be linear, let $\mathcal{B}, \mathcal{B}'$ be bases for V . Given $[L]_{\mathcal{B}, \mathcal{B}}$, how can one compute $[L]_{\mathcal{B}', \mathcal{B}'}$?

One can answer 1. question with the **transition matrix** or **change of basis matrix**

$$[v]_{\mathcal{B}'} = [id]_{\mathcal{B}', \mathcal{B}} \cdot [v]_{\mathcal{B}} \quad (22)$$

One can answer 2. question similarly. For an invertible matrix P that goes from \mathcal{B} to \mathcal{B}' , we have:

$$[L]_{\mathcal{B}', \mathcal{B}'} = P^{-1} [L]_{\mathcal{B}, \mathcal{B}} P \quad (23)$$

This leads us to the notion of similar matrices.

Definition 17 (Similar). We say that two $n \times n$ matrices are **similar** there exists an invertible matrix $P \in \text{MAT}(n \times n, \mathbb{K})$ such that

$$B = P^{-1} A P \quad (24)$$

3.7 The Normal Form Problem

Definition 18 (Equivalence Relation). Let S be a set. A **relation** is a subset of the cartesian product $S \times S$. An equivalence relation is denoted by \sim , and it holds for any relation such that the following properties hold:

1. Every element in S is in a relation with itself:

$$\forall x \in S, x \sim x \quad (25)$$

2. (Reflexivity): $\forall x, y \in S, x \sim y \iff y \sim x$.
3. (Transitivity): $\forall x, y, z \in S$ if $x \sim y \wedge y \sim z \Rightarrow x \sim z$.

Equivalence relations partitions sets S into pairwise disjoint equivalence classes.

Definition 19 (Equivalence Class). Let S be a set. Let \sim be an equivalence relation on S . Then, the **equivalence class** of S is defined as:

$$[s]_{\sim} := \{t \in S \mid t \sim s\} \quad (26)$$

where $s \in S$.

Theorem 35. Let S be a set, \sim an equivalence relation on S . The equivalence classes of \sim partition S (i.e, every $s \in S$ is contained in an equivalence class) and any two equivalence classes are either identical or disjoint).

Theorem 36. Similarity is an equivalence relation on the set of all $n \times n$ matrices with coefficients in K . That is, the relation

$$B \sim A \iff \exists P \in GL(n, K) \text{ such that } B = P^{-1} A P \quad (27)$$

is an equivalence relationship.

Motivating Question of the Normal Form Problem: can we within a given similarity class find one representation or several that is “simpler” than the others? These can be more useful, since they can shed light on certain mathematical details. This is the motivation behind eigenvectors and eigenvalues.

Definition 20 (Invariant Subspace). Let $L : V \rightarrow V$ be a linear map. Let V be a vector space over K , can be finite or infinite-dimensional. A subspace $U \leq V$ is called **L-Invariant** if

$$\text{im}[L(u)] \subseteq U \tag{28}$$

for example, the kernel of any linear map is invariant since its mapped to zero, and $0 \in U$.

Definition 21 (Eigenvector/Eigenvalue). Let $(V, +, \cdot)$ be a vector space over K . Let $L : V \rightarrow V$ be linear. Let $v \in V$, $v \neq 0$, such that $L(v) = \lambda v$ for some $\lambda \in K$. Then, v is called a **eigenvector** of L to the **eigenvalue** λ .

Theorem 37. Let V be a finite-dimensional vector space over K . Let $L : V \rightarrow V$ be a linear map. Let λ be an eigenvalue of L . Define $E_\lambda := \{v \in V \mid L(v) = \lambda v\}$. Then, $E_\lambda(L) \leq V$. This is called the **eigenspace** of the eigenvalue λ .

Theorem 38. $E_\lambda(L)$ is L -invariant.

Theorem 39. (Eigenvectors to distinct eigenvalues are linearly dependent). If $L : V \rightarrow V$ is linear, v_1, \dots, v_k are eigenvectors of L to eigenvalues $\lambda_1, \dots, \lambda_k$ (which are pairwise distinct), then v_1, \dots, v_k are linearly independent.

Theorem 40 (Very Fundamental and Important). Let V be a vector space over K , V finite-dimensional. Let L be an endomorphism. Then, L is diagonalisable, i.e., there is a basis \mathcal{B} for which $[K]_{\mathcal{B}}$ is a diagonal matrix $\iff V$ has a basis of eigenvectors.

Question: can we find a criteria for a linear map to be diagonalisable?

Theorem 41. Let V be a vector space over K with dimension of V equal to n . Let $L : V \rightarrow V$ be a linear map. Assume that L has a basis of eigenvectors to pairwise distinct eigenvalues. Then, L is diagonalisable.

Question: is every linear map $L : V \rightarrow V$ diagonalisable?

Answer: No. Consider the following examples:

Example 27. • Differentiation map \mathcal{D} is not diagonalisable.

Theorem 42. Let $L : V \rightarrow V$. Then, 0 is an eigenvalue of $L \iff L$ is not injective. In this case, the eigenspace to the eigenvalue 0 is the kernel.

3.8 Characteristic Polynomial

Theorem 43. Similar matrices have the same characteristic polynomial.

3.9 Normal Form for Non-Diagonalisable Matrices

Definition 22 (Upper/Lower Triangular). Let $A \in \text{MAT}(n \times n, \mathbb{K})$. A matrix of the form:

$$A = \begin{bmatrix} \lambda_1 & * & * & * & \dots & * \\ 0 & \lambda_2 & * & * & \dots & * \\ 0 & 0 & \lambda_3 & * & \dots & * \\ 0 & 0 & \dots & \ddots & \dots & * \\ 0 & 0 & 0 & 0 & 0 & \lambda_n \end{bmatrix} \tag{29}$$

is called **upper triangular**. The transpose of such a matrix is called **lower triangular**. Here, $\lambda_1, \dots, \lambda_n$ are the eigenvalues.

Theorem 44. Let $A \in \text{MAT}(n \times n, \mathbb{C})$ and let A be upper triangular. Then, the elements on the main diagonal are exactly the eigenvalues of A .

Theorem 45 (“Great Theorem”). Every matrix in $\text{MAT}(n \times n, \mathbb{C})$ is similar to an upper triangular matrix.

4 Applications of Linear Algebra (to other fields of math...not real life :-P)

4.1 Matrix Exponential

Recall from Math 222 the MacLaurin series of e :

$$e^x = 1 + x + \frac{x^2}{2!} + \dots \quad (30)$$

$$= \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (31)$$

Can we do the same with matrices? That is, let $X \in \text{MAT}(n \times n, \mathbb{K})$. Does the limit

$$\lim_{N \rightarrow \infty} \sum_{k=0}^N \frac{1}{k!} X^k \quad (32)$$

exist in the analytic sense? Yes – w/o justification, see analysis. The limit is denoted by:

$$\exp(X) := \sum_{k=0}^{\infty} \frac{1}{k!} X^k \quad (33)$$

this is called the **matrix exponential of X** . This is used to solve ODEs. In short, the matrix exponential of a diagonal matrix is really nice. One can diagonalise a non-diagonal matrix and take the exponential. If the matrix is not diagonalisable, then one can upper-triangularise it. Then, the matrix exponential gives you a recursive relation which you can then use a computer to solve.

This can also be used to solve difference equations.

5 Orthogonality: A quick introduction

5.1 Standard inner product on \mathbb{R}^n

Recall the standard inner product on \mathbb{R}^n . Let $u = [u_1, \dots, u_n]^t$ and $v = [v_1, \dots, v_n]^t$. Then:

$$\langle u, v \rangle := u_1 v_1 + \dots + u_n v_n \quad (34)$$

This is called a **bilinear form**, i.e., it’s linear in each component. What does this mean?

$$\langle u + w, v \rangle = \langle u, v \rangle + \langle w, v \rangle \quad (35)$$

$$\langle ku, v \rangle = k \langle u, v \rangle \quad (36)$$

$$\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle \quad (37)$$

$$\langle u, kv \rangle = k \langle u, v \rangle \quad (38)$$

Now let’s go to \mathbb{C} . Let $z = x + iy$. Recall the **complex conjugate** of z , \bar{z} :

$$\bar{z} = \overline{x + iy} = x - iy \quad (39)$$

We have the following:

$$\bar{z}z = |z|^2 \quad (40)$$

Let's generalise this to \mathbb{C}^n . How can we determine the length of a vector in \mathbb{C}^n ? To do this, first define:

$$z^* := \bar{z}^t \quad (41)$$

then,

$$\|z\|^2 = z^*z \quad (42)$$

We will thus define the standard inner product on \mathbb{C}^n as:

$$\langle z, w \rangle := z^*w = \bar{z}_1w_1 + \dots + \bar{z}_nw_n \quad (43)$$

5.2 Properties of the Standard Inner Product on \mathbb{C}^n

Caution! The standard inner product on \mathbb{C}^n is not bilinear; it is only sesqui-linear. Here, it is only linear in its FIRST component:

$$\langle z, v + w \rangle = \langle z, v \rangle + \langle z, w \rangle \quad (44)$$

$$\langle z, kw \rangle = k\langle z, w \rangle \quad (45)$$

$$\langle z + v, w \rangle = \langle z, w \rangle + \langle v, w \rangle \quad (46)$$

$$\langle kz, w \rangle = \bar{k}\langle z, w \rangle \quad (47)$$

Definition 23 (Orthogonal). Two vectors $z, w \in \mathbb{C}^n$ are called orthogonal if their inner product $\langle z, w \rangle = 0$.

Definition 24 (Unit Vector). A vector $v \in K^n$ is a unit vector if $\|v\| = \sqrt{\langle v, v \rangle} = 1$.

Definition 25. Vectors $\{v_1, \dots, v_n\} \in K^n$ are called orthogonal if $\{v_1, \dots, v_n\}$ are pairwise orthogonal. They're called orthonormal if they're pairwise orthogonal and all the v_i are unit vectors.

If $\{v_1, \dots, v_k\}$ are orthonormal and $v = a_1v_1 + \dots = a_nv_n$, then

$$a_j = \langle v_i, v \rangle \quad (48)$$

Definition 26 (Orthogonal). A matrix $A \in \text{MAT}(n \times n, \mathbb{R})$ is called orthogonal if the columns of A are orthonormal ("for historical reasons – a bit of an unfortunate thing").

Definition 27 (Unitary). A matrix $A \in \text{MAT}(n \times n, \mathbb{C})$ is called unitary if the columns of A are orthonormal with respect to the inner product on \mathbb{C}^n .

5.3 Application of Orthogonality (Orthogonal and Unitary Matrices)

5.3.1 Orthogonal Case

Let $A \in \text{MAT}(n \times n, \mathbb{R})$ be orthogonal. Then, its inverse is A^t .

5.3.2 Unitary Case

Let $A \in \text{MAT}(n \times n, \mathbb{C})$ be unitary. Then, its inverse is A^* , where $A^* := \bar{A}^t$.

Question: let $U \leq K^n$. Let $\mathcal{B} = \{v_1, \dots, v_k\}$ be a basis for U . How can we find an orthogonal or maybe orthonormal basis \mathcal{B}' for the same subspace?

Answer: Gram-Schmidt process !! Let $\mathcal{B} = \{v_1, \dots, v_k\}$ be a basis for U . Then, define the following vectors:

$$\begin{aligned} u_1 &= v_1 \\ u_2 &= v_2 - \frac{\langle u_1, v_2 \rangle}{\langle u_1, u_1 \rangle} u_1 \\ u_3 &= v_3 - \frac{\langle u_1, v_3 \rangle}{\langle u_1, u_1 \rangle} u_1 - \frac{\langle u_2, v_3 \rangle}{\langle u_2, u_2 \rangle} u_2 \\ &\vdots \\ u_k &= v_k - \frac{\langle u_1, v_k \rangle}{\langle u_1, u_1 \rangle} u_1 - \dots - \frac{\langle u_{k-1}, v_k \rangle}{\langle u_{k-1}, u_{k-1} \rangle} u_{k-1} \end{aligned}$$

When the GS process terminates, make sure you normalise each vector to get an orthonormal basis (as of now, you just have an orthogonal basis).

Theorem 46. Let P be orthogonal. Then, the row vectors of P are pairwise orthogonal unit vectors.

Theorem 47. 1. Let P, Q be orthogonal. Then, P^{-1} is orthogonal and the product PQ is orthogonal. The set of all orthogonal matrices form a subgroup of general linear with respect to matrix multiplication:

$$O(n) := \{P \in \text{MAT}(n \times n, \mathbb{R}) \mid P \text{ orthogonal} \} \quad (49)$$

2. Let P, Q be unitary. Then P^{-1} is unitary and PQ is unitary. Moreover:

$$U(n) := \{P \in \text{MAT}(n \times n, \mathbb{C}) \mid P \text{ unitary} \} \quad (50)$$

is a group with respect to matrix multiplication, and is in fact a subgroup of $GL(n, \mathbb{C})$.

5.4 Orthogonality and Unitary: Diagonalisation and Upper Triangularisation

Question: what matrices can be diagonalised or upper triangularised by an orthogonal matrix? What matrix can be diagonalised or upper triangularised by a unitary matrix?

Theorem 48 (Schur). Every matrix $A \in \text{MAT}(n \times n, \mathbb{C})$ can be upper-triangularised by a unitary matrix.

Corollary 1. If $A \in \text{MAT}(n \times n, \mathbb{R})$ has only real eigenvalues, A can be upper-triangularised via an orthogonal matrix.

Theorem 49. Let $A \in \text{Symm}(n \times n, \mathbb{R})$.

1. Then, all eigenvalues of A are real.
2. Eigenvectors to distinct eigenvalues of A are automatically orthogonal.

Theorem 50. Let $A \in \text{Symm}(n \times n, \mathbb{R})$. and let $P \in O(n)$. Then, $B = P^{-1}AP = P^tAP$ is again symmetric.

Theorem 51 (Spectral Theorem in \mathbb{R}). A matrix $A \in \text{MAT}(n \times n, \mathbb{R})$ is diagonalisable via an orthogonal matrix $P \iff A$ is symmetric.

5.4.1 Spectral Theorem: Complex Case

Question: what types of matrices can be diagonalised via a unitary matrix?

Definition 28 (Hermitian). A matrix $A \in \text{MAT}(n \times n, \mathbb{C})$ is called hermitian if $A^* = A$ (complex equivalent of a real symmetric matrix).

One may conjecture that A is diagonalisable by a unitary matrix if A is hermitian. However, while it is true that all Hermitian matrices are diagonalisable via a unitary matrix, the converse does not hold. We need some more definitions.

Definition 29 (Normal Matrix). A matrix $A \in \text{MAT}(n \times n, \mathbb{C})$ is called **normal** if $AA^* = A^*A$.

Many important matrices are normal.

Theorem 52. The following classes of matrices are normal:

1. Real symmetric matrices.
2. Skew-symmetric real matrices (all eigenvalues are imaginary).
3. Real orthogonal matrices (here, the determinant's absolute value is one; resembles rotations and reflections).
4. Hermitian matrices
5. Skew-Hermitian matrices ($A^* = -A$).
6. Unitary matrices.

Remark. If A is normal, $P \in GL(n, \mathbb{C})$, then $P^{-1}AP$, in general, is not normal. However...

Theorem 53. Let $A \in \text{MAT}(n \times n, \mathbb{C})$, $P \in U(n)$. Then, $P^{-1}AP = P^*AP$ is normal.

Theorem 54. Let $A \in \text{MAT}(n \times n, \mathbb{C})$ be both normal and upper triangular. Then, A is a diagonal matrix.

Theorem 55 (Complex Spectral Theorem). Let $A \in \text{MAT}(n \times n, \mathbb{C})$. Then, A is diagonalisable by a unitary matrix $\iff A$ is normal.

Definition 30. A real symmetric matrix is called...

1. **Positive definite** if all eigenvalues of A are positive.
2. **Positive semi-definite** if all eigenvalues are greater or equal than zero.
3. **Negative definite** if all eigenvalues are strictly less than zero.
4. **Negative semidefinite** if all eigenvalues are less than or equal to zero.

5.5 Applications of the Real Spectral Theorem

5.5.1 Real Quadratic Forms

Definition 31 (Real Quadratic Form). A **real quadratic form** is a purely quadratic polynomial in n variables with real coefficients. For example:

$$X_1^2 + 4X_1X_2 + X_2^2 \tag{51}$$

Question: How can we determine the range of a quadratic form?

Question: Can every quadratic form be written as a sum or difference of squares?

Definition 32 (Positive Definite). A quadratic form $X := Q(X_1, \dots, X_n)$ is called **positive definite** if $Q(x) \geq 0 \forall X \in \mathbb{R}^n$ and $Q(x) = 0 \iff X = 0$. One can equivalently define negative definite.

Every quadratic form $Q(X)$ can be written in “matrix form”, i.e;

$$Q(X) = X^tAX \tag{52}$$

6 Inner Product Spaces

There is no notion of length in an abstract vector space. Therefore, the aim is to introduce a notion of lengths and angles to abstract vector spaces, i.e., generalise what we've done with the standard inner product.

We generalise between $K = \mathbb{R}$ and $K = \mathbb{C}$.

6.1 $K = \mathbb{R}$

Definition 33 (Inner Product Space). Let V be a vector space over \mathbb{R} . A function: $\langle, \rangle : V \times V \rightarrow \mathbb{R}$ is called an inner product on V if:

1. \langle, \rangle is bilinear
2. \langle, \rangle is symmetric
3. \langle, \rangle is positive definite

A vector space with an inner product is called a inner product space.

Example 28. The standard inner product on \mathbb{R}^n , $\langle x, y \rangle := x^t y$ is an inner product.

Example 29. Let $V = \mathbb{R}^n$ a positive definite $n \times n$ (which implies symmetric on \mathbb{R} or hermitian on \mathbb{C}) matrix A gives us an inner product:

$$\langle x, y \rangle := x^t A y \quad (53)$$

Example 30. Let $V = C[a, b]$: the vector space of all continuous functions on $[a, b]$. A useful inner product in analysis is:

$$\langle f, g \rangle := \int_a^b f(x)g(x)dx \quad (54)$$

6.2 Complex Inner Product Spaces

Definition 34 (Complex Inner Product Space). Let V be a VS over \mathbb{C} , $\langle, \rangle : V \times V \rightarrow \mathbb{C}$ such that

1. \langle, \rangle is sesqui-linear.
2. Conjugate symmetry: $\forall z, w \in V$:

$$\langle z, w \rangle = \overline{\langle w, z \rangle} \quad (55)$$

3. Positive definite.

Then, (V, \langle, \rangle) is called a complex inner product space.

Example 31. Let $A \in \text{MAT}(n \times n, \mathbb{C})$ be a Hermitian positive definite matrix. Then $\langle z, w \rangle := z^* A w$ is a complex inner product on \mathbb{C}^n .

6.3 Geometry on a Vector Space

We can define the length on an IPS as:

$$\|x\| := \sqrt{\langle x, x \rangle} \quad (56)$$

Claim: this satisfies all of the conditions in what mathematics is called a norm.

Definition 35 (Norm). Let $(V, +, \cdot)$ be a vector space over K . Then, $\|\cdot\|$ (norm) is a function that assigns a length to a vector: $V \rightarrow \mathbb{R}_0^+ = [0, \infty[$ such that

1. $\|x\| \geq 0 \forall x \in V$ and $\|x\| = 0 \iff x = 0$.
2. $\|kx\| = |k|\|x\| \forall k \in K, x \in V$.
3. Triangle Inequality: $\|x + y\| \leq \|x\| + \|y\|$.

Theorem 56. Let V be an IPS over K , $\|x\| := \sqrt{\langle x, x \rangle} \forall x \in V$. Then, $\|\cdot\|$ is a norm on V .

Theorem 57 (Cauchy-Schwarz Inequality). Let V be an IPS over K . Then:

$$|\langle x, y \rangle| \leq \|x\|\|y\| \quad (57)$$

$\forall x, y \in V$.

Lemma 58. $\langle a + \lambda b, a + \lambda b \rangle = \langle a, a \rangle + \lambda \langle a, b \rangle + \overline{\lambda} \overline{\langle a, b \rangle} + \lambda \overline{\lambda} \langle b, b \rangle$

Theorem 59 (Triangle Inequality). Let $x, y \in V$ where V is an IPS. Then:

$$\|x + y\| \leq \|x\| + \|y\| \quad (58)$$

(this is the most important theorem in maths).

Lemma 60. Let $z \in \mathbb{C}$. Then:

1. $z + \bar{z} = 2 \operatorname{Re}(z)$
2. $|\operatorname{Re}(z)| \leq |z|$

Definition 36 (Angle). Let V be a real IPS with an inner product \langle, \rangle . Let $x, y \neq 0$ in V . Let $\varphi \in [0, \pi]$ be uniquely determined such that

$$\cos(\varphi) = \frac{\langle x, y \rangle}{\|x\| \|y\|} \in [-1, 1] \quad (59)$$

(where the inclusion is thanks to Cauchy Schwarz). Then φ is the angle between x and y .

Definition 37. Let V be a real IPS. We say that x and y are orthogonal if $\langle x, y \rangle = 0$. Note that for $x, y \neq 0$, this is equivalent to the angle between x and y being $\pi/2$.

Definition 38. Let V be an IPS over \mathbb{R} . We say that $x, y \in V$ are orthogonal \iff their inner product is zero.

Theorem 61 (Pythagorean's Theorem). Let V be an IPS. Let $X, Y \in V$ be orthogonal. Then

$$\|X + Y\|^2 = \|X\|^2 + \|Y\|^2 \quad (60)$$

7 Fourier Series in 30 Minutes [not examinable]

Consider V the vector space of all continuous functions on \mathbb{R} endowed with the inner product

$$\int_{-\pi}^{\pi} f(x)g(x)dx \quad (61)$$

and consider the following set

$$S := \{1, \sin(x), \cos(x), \sin(2x), \cos(2x), \dots\} \quad (62)$$

Claim: the functions in S are pairwise orthogonal:

$$\int_{-\pi}^{\pi} \sin(nx) \cos(mx) dx = 0 \quad (63)$$

$\forall n \in \mathbb{N}, m \in \mathbb{N}_0$. This follows from the product-to-sum trigonometric identities. Moreover, observe:

$$\int_{-\pi}^{\pi} \sin^2(nx) dx = \pi \quad \forall n \in \mathbb{N} \quad (64)$$

$$\int_{-\pi}^{\pi} \cos^2(nx) dx = \pi \quad \forall n \in \mathbb{N} \quad (65)$$

$$\int_{-\pi}^{\pi} 1 dx = 2\pi \quad \text{for } n = 0 \quad (66)$$

Now recall the following: Let's work in \mathbb{R}^n with the standard inner product. Let v_1, \dots, v_n be pairwise orthogonal, non-zero vectors. Then, we may write:

$$v = a_1 v_1 + \dots + a_n v_n \quad (67)$$

for any vector $v \in \mathbb{R}^n$. Recall that we recovered the coefficients a_1, \dots, a_j previously with the formula:

$$a_j = \frac{\langle v_j, v \rangle}{\langle v_j, v_j \rangle} \quad (68)$$

(the proof for that only used linearity in the second argument and positive definiteness). Now let $f(x)$ represent a “signal”:

$$f(x) = c_0 + [a_1 \cos(x) + \dots + a_n \cos(nx)] + [b_1 \sin(x) + \dots + b_n \sin(nx)] \quad (69)$$

We know that $f(x)$ is a linear combination of the members in S . Assume that we are given an $f(x)$. **How can we recover the linear combination of sines and cosines?** Well, we simply apply the formula:

$$a_n = \frac{\langle \cos(nx), f(x) \rangle}{\langle \cos(nx), \cos(nx) \rangle} \quad (70)$$

$$= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx \quad (71)$$

This is all linear algebra! No calculus or analysis! Similarly for the b_n 's:

$$b_n = \frac{\langle \sin(nx), f(x) \rangle}{\langle \sin(nx), \sin(nx) \rangle} \quad (72)$$

$$= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx \quad (73)$$

and we can recover c_0 :

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx \quad (74)$$

Drawback: in linear algebra, all sums must be finite. Not all functions can be represented as a finite sum of trigonometric functions; they must also be periodic.

Question: Can $f(x)$ be represented as an infinite series? Does it converge?

We will use these formulae as a motivation for the definition of a Fourier series.

Definition 39 (Fourier Series). Let $f(x)$ be periodic and continuous with period 2π . Define:

$$a_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx \quad (75)$$

$$b_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx \quad (76)$$

$$c_0 := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx \quad (77)$$

Now, the analysis comes; two questions:

1. Does it converge?
2. Will we get $f(x)$ back?

If it does converge, we obtain the **Fourier Series** of f , which is entirely generated by orthogonality relations. There is a lot of substantial literature covering these two questions.

Theorem 62 (Dirichlet's Theorem). Let F be a periodic function with period 2π and let it be piecewise continuous and continuously differentiable on $[-\pi, \pi]$. Then, the Fourier series of f converges at all points where f is continuous, and it represents f .

This is pretty good news as long as your functions are reasonably well-behaved.