MAS277 Vector Spaces and Fourier Theory

Semester 2, 2013–14

Course details

Lecturer details:

- Dr. Fionntan Roukema
- f.roukema@sheffield.ac.uk
- K12 Hicks building.

Lectures: Lectures take place twice a week in weeks 1-11.

- Monday 10:00 Hicks LT2
- Thursdays 09:00 Hicks LT7.

Tutorials: Tutorials take place in weeks 1, 3, 5, 7, 9, 11.

- Group 1: Tuesdays 11:00 Hicks F20,
- Group 2: Tuesdays 14:00 Hicks F20,
- Group 3: Tuesdays 16:00 Hicks F24.

Course material: All material for this course will appear on MOLE as well as the course webpage http://roukema.staff.shef.ac.uk/mas277.html. Course material will include:

- Skeleton notes.
- Worked Example booklet.
- (subject to willing volunteers) scanned versions of student notes.
- Problem booklet.
- Past and sample exam papers.

Assessment:

- Your mark will be completely determined by the result of a 2 hour exam that will count for 100% of your final mark. The exam will take place during the Summer exam period.
- To give you feedback as the course progresses there will be 4 assignments that will be submitted and returned during your tutorials. These assignments will not contribute to your final mark.

Additional help: In addition to lectures and tutorials:

- I always welcome questions and discussion and you should speak to me after lectures and tutorials whenever you have questions.
- Official office hours will take place on Wednesday morning.
- You should also feel welcome to speak to me outside office hours (to avoid a wasted trip, ring 0114 222 3872 to make sure I am in/available).

What is MAS277 and why/how should I engage with it?

The subject matter. MAS277 is a first course in linear algebra from an abstract view point. In particular, we will abstract the "essential qualities" that $\mathbb{R}^n$ possesses and take them to be the axioms of more general objects. This generalised approach has two main advantages.

Firstly, theorems are now true in any setting that satisfies our axioms. So you can prove infinitely many theorems for the price of one! For example, we will see that solving problems like finding

$$(*) \qquad \min\left\{\int_a^b (f(x) + a_n x^n + \cdots + a_0)^2 \mathrm{dx} : \mathrm{a_i} \in \mathbb{R}\right\}$$

for a fixed continuous function $f(x)$ can be solved in "exactly" the same way as computing the distance from a point to a plane in $\mathbb{R}^3$. We will see that Fourier series can be viewed in the same way as Problem $(*)$. That is, finding Fourier polynomials is like finding a point in a given plane that is closest to a given point.

Secondly, proofs of theorems in the abstract setting are often much easier. We will see a trivial one line proof of Pythagoras' theorem which is true in every "reasonable" setting (including infinite dimensional spaces).

What is expected of me.

- You are expected to attend all lectures and tutorials. You are also expected to speak to your tutors and lecturer about mathematics.
- Lectures will aim to be as engaging and interactive as possible. As a result, examples that are presented in lectures are likely to differ from those suggested in the notes. So, any examples in the notes that are not fully worked through in lectures should be viewed as additional problems for your problem booklet.
- Mathematics is an active sport and is best learned by solving problems. You are expected to work through as many of the examples from your problem booklet/lectures/worked example booklet/lecture notes/recommended texts as you can.
- You are expected to submit all homework.

What will I need to do if I want to do well on this course.

- You will need to work through as many of the examples from your problem booklet/lectures/worked example booklet/lecture notes/recommended texts as you can.
- You will need to be familiar with all notation in this course.
- You will need to be able to carefully state all definitions and major results from

this course.
- You will need to be able to prove when an example is or is not a defined quantity from this course (e.g. be able to show if something is a vector space etc.).
- You will need to be able to apply the ideas and techniques from this module to obtain proofs of elementary results.

CHAPTER 1

Vector spaces

1. Fields

Vectors in $\mathbb{R}^n$ may be scaled by any real number. Real numbers can be added and subtracted as well as multiplied and divided. We start with an abstraction of a set in which you can add, subtract, multiply and divide.

Definition 1. A *field* $\mathbb{F}$ is a set together with two binary operations[1] "+" and "·" with the following properties:

F1: $\mathbb{F}$ is an Abelian group[2] with respect to the binary operation "+".
F2: $\mathbb{F}\backslash\{e_+\}$[3] is an Abelian group with respect to the binary operation "·".
F3: Multiplication is distributive with respect to addition. Namely, for every $a, b, c \in \mathbb{F}$ we have

$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

Remark 1. Both **F1** and **F2** can each be unpackaged into 4 axioms; associativity, existence of identity, existence of inverses, commutativity (see Footnote 2). So, a field is defined by 9 axioms.

Remark 2. To draw a closer analogy with $\mathbb{R}$ we adopt the following conventions:

- For $a, b \in \mathbb{F}$ the element $+(a, b) \in \mathbb{F}$ is written $a + b$.
- The identity with respect to $+$ is denoted by 0 and pronounced "zero".
- The inverse of $a \in \mathbb{F}$ is denoted $-a$. We write $a - b$ as a shorthand for $a + (-b)$.
- For $a, b \in \mathbb{F}$ the element $\cdot(a, b) \in \mathbb{F}$ is written $a \cdot b$ or ab.

---

[1] A binary operation on a set X is a map $* : X \times X \to X$

[2] an Abelian group is a set G together with a binary operation $*$ that satisfies the following properties: (associativity) $\forall g_1, g_2, g_3 \in G, (g_1 * g_2) * g_3 = g_1 * (g_2 * g_3)$, (existence of identity) $\exists e \in G$ such that $\forall g \in G, e * g = g * e = g$, (existence of inverses) $\forall g \in G, \exists h \in G$ s.t. $g * h = h * g = e$, (commutativity) $\forall g_1, g_2 \in G, g_1 * g_2 = g_2 * g_1$

[3] Here e_+ denotes the identity of the group $\mathbb{F}$ with respect to "+"

- The identity of $\mathbb{F}\backslash\{0\}$ with respect to $\cdot$ is denoted by 1 and pronounced "one".
- The inverse of $a \in \mathbb{F}$ is denoted by a^{-1} or $\frac{1}{a}$. We will sometimes write $\frac{a}{b}$ as a shorthand for ab^{-1}.

The following proposition shows us that **F1-F3** guarantee that fields possess most reasonable properties that you would expect an abstraction of $\mathbb{R}$ to possess.

PROPOSITION 1.1. *Let a, b, c be arbitrary elements of a field $\mathbb{F}$. The following statements are true:*
(a) If $a + b = c + b$ then $a = c$.
(b) If $a \cdot b = c \cdot b$ and $b \neq 0$ then $a = c$.
(c) $a \cdot 0 = 0$.
(d) $(-a) \cdot b = a \cdot (-b) = -(a \cdot b)$.
(e) $(-a) \cdot (-b) = a \cdot b$.

Example 1. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are all fields with respect to the usual notions of addition and multiplication.

Example 2. For a positive prime integer p, the quotient $\mathbb{Z}/p\mathbb{Z}$ is a field with respect to the usual notions of addition and multiplication.

2. Abstract vector spaces

We begin with the definition of a vector space.

Definition 2. A *vector space* V over a field $\mathbb{F}$ is a set with a binary operation $+ : V \times V \to V$ (called *addition*) and a function $\cdot : \mathbb{F} \times V \to V$ (called *scalar multiplication*) such that Properties **VS1-VS5** hold.

VS1: V is an Abelian group with respect to $+$.
VS2: For every $\mathbf{v} \in V$ we have $1\mathbf{v} = \mathbf{v}$.
VS3: For all $a, b \in \mathbb{F}$ and $\mathbf{v} \in V$ we have $(ab)\mathbf{v} = a(b\mathbf{v})$.
VS4: For all $a \in \mathbb{F}$ and $\mathbf{u}, \mathbf{v} \in V$ we have $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$.
VS5: For all $a, b \in \mathbb{F}$ and $\mathbf{v} \in V$ we have $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$.

The elements of V are called *vectors* and the elements of $\mathbb{F}$ are called *scalars*.

Remark 3. In Definition 2 we used the conventions that $+(\mathbf{u}, \mathbf{v})$ was written as $\mathbf{u} + \mathbf{v}$ and $\cdot(a, \mathbf{v})$ was written as $a\mathbf{v}$. We will use the conventions that the identity of V with respect to addition is denoted $\mathbf{0}$ and that the inverse of $\mathbf{v} \in V$ with respect to $+$ is written as $-\mathbf{v}$.

Remark 4. The axiom **VS1** can be unpackaged into 4 axioms (associativity, existence of identity, existence of inverses, and commutativity). So, a vector space is defined by 8 axioms.

PROPOSITION 1.2. *The following statements are true for any vector space V over a field $\mathbb{F}$:*
(a) $0\mathbf{v} = \mathbf{0}$ *for every* $\mathbf{v} \in V$.
(b) $(-a)\mathbf{v} = -(a\mathbf{v}) = a(-\mathbf{v})$ *for every* $a \in \mathbb{F}$ *and* $\mathbf{v} \in V$.
(c) $a\mathbf{0} = \mathbf{0}$ *for every* $a \in \mathbb{F}$.

Example 3. We now give a long list of common examples of vector spaces.

(1) $\mathbb{R}^n$ over $\mathbb{R}$.
(2) $\mathbb{C}^n$ over $\mathbb{C}$.
(3) $\mathbb{R}[x] := \{p(x) : p(x) \text{ is a polynomial with coefficients in } \mathbb{R}\}$ over $\mathbb{R}$ with addition and scalar multiplication defined in the obvious way.
(4) $\mathbb{R}[x]_{\leq n} := \{p(x) \in \mathbb{R}[x] : \deg(p(x)) \leq n\}$ over $\mathbb{R}$ with addition and scalar multiplication defined as for $\mathbb{R}[x]$.
(5) Let X and Y be vector spaces over a field $\mathbb{F}$. The space
$$\mathcal{L}(X, Y) := \{f : X \to Y : f(x_1 + x_2) = f(x_1) + f(x_2)\}$$
over $\mathbb{F}$ with addition defined by $(f + g)(x) = f(x) + g(x)$ and scalar multiplication defined by $(cf)(x) = c(f(x))$ is a vector space.
(6) For X a set, $\mathcal{F}(X, \mathbb{F}) := \{\text{functions } f : X \to \mathbb{F}\}$ over $\mathbb{F}$ with addition and scalar multiplication defined as for $\mathcal{L}(X, \mathbb{F})$.
(7) $C[a, b] := \{f : [a, b] \to \mathbb{R} : f \text{ is continuous}\}$ over $\mathbb{R}$ with addition and scalar multiplication defined in the obvious way.
(8) $C^\infty[a, b] := \{f : [a, b] \to \mathbb{R} : f \text{ is smooth}\}$ over $\mathbb{R}$ with addition and scalar multiplication defined in the obvious way.
(9) The set $\mathcal{S}(\mathbb{F})$ of sequences of real numbers $\{a_n\}_{n \in \mathbb{N}}$ over $\mathbb{F}$ with addition defined by $\{a_n\} + \{b_n\} = \{a_n + b_n\}$ and scalar multiplication defined by $c\{a_n\} := \{ca_n\}$.
(10) The set $M_{n \times m}(\mathbb{F})$ of $n \times m$ matrices with enteries in $\mathbb{F}$ with the usual notions of matrix addition and scalar multiplication.

Remark 5. We will often not specify the field $\mathbb{F}$ when the context is clear or not important. So, "let V be a vector space" means "let V be a vector space over a field $\mathbb{F}$".

Definition 3. Let V be a vector space over a field $\mathbb{F}$. A set $W \subseteq V$ is called a *subspace* of V if the following conditions hold:

- $\mathbf{0} \in W$.
- $\mathbf{u} + \mathbf{v} \in W$ whenever $\mathbf{u}, \mathbf{v} \in W$.
- $c\mathbf{v} \in W$ for all $\mathbf{v} \in W$ and $c \in \mathbb{F}$.

Example 4. Let V be a vector space. Both $\{\mathbf{0}\}$ and V are subspaces of V.

Example 5. Taking the vector spaces defined in Example 3 we can see that (4) is a subspace of (3), and (8) is a subspace of (7), and both (7) and (8) are subspaces of $\mathcal{F}([a, b], \mathbb{R})$.

PROPOSITION 1.3. *Let W be a subspace of a vector space V over a field $\mathbb{F}$. The following statements are true:*

(a) If W is a subspace of V then W is a vector space over $\mathbb{F}$.
(b) $U \subseteq V$ is a subspace if and only if
- $\mathbf{0} \in U$
- $\mathbf{u} + c\mathbf{v} \in U$ *whenever* $\mathbf{u}, \mathbf{v} \in U$ *and* $c \in \mathbb{F}$.

(c) If $\{U_i\}_{i \in I}$ is a collection of subspaces of V then $\cap_{i \in I} U_i$ is a subspace of V.
(d) If U and W are subspaces of V then

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

is a subspace of V.

Definition 4. The space $U + V$ defined in Proposition 1.3 is called *the sum* of U and V. When $U \cap V = \{\mathbf{0}\}$ then the sum of U and V is called *the direct sum* and is written $U \oplus V$.

3. Bases and dimension

From MAS201 we are already familiar with the standard basis vector $\mathbf{e}_i \in \mathbb{R}^n$ whose only non-zero coordinate is the i^{th} coordinate which is 1. For example, $\mathbf{e}_4 \in \mathbb{R}^5$ is given by

$$\mathbf{e}_4 = (0, 0, 0, 1, 0)^T.$$

We also know that all $\mathbf{x} \in \mathbb{R}^n$ can be written uniquely as a linear sum of vectors in $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. We now work towards defining the natural abstraction of a basis in the setting of an abstract vector space.

Definition 5. Let V be a vector space over a field $\mathbb{F}$. A set $U \subseteq V$ is said to be *linearly dependent* if there exists $\mathbf{u}_1, \ldots, \mathbf{u}_n \in U$ and $c_1, \ldots, c_n \in \mathbb{F}\backslash\{0\}$ such that $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{0}$. A set $W \subseteq V$ which is not linearly dependent is said to be *linearly independent.*

Example 6.

(a) If $U \subset V$ and $\mathbf{0} \in U$ then U is linearly dependent.
(b) The set $\{(1, 1, 1)^T, (1, -1, -1)^T, (1, 0, 0)^T\} \subset \mathbb{R}^n$ is linearly dependent.
(c) The set $\{1, x, \ldots, x^n\} \subset \mathbb{R}[x]$ is linearly independent.
(d) A subset of a linearly independent set is linearly independent.

Definition 6. Let V be a vector space over a field $\mathbb{F}$, and let $U \subseteq V$. We define

$$\mathrm{sp}(U) := \{\mathbf{u} \in V : \text{ there exists } \mathbf{u}_1, \ldots \mathbf{u}_k \in U, c_1, \ldots, c_k \in \mathbb{F} \text{ with } \mathbf{u} = c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k\}.$$

The set $\mathrm{sp}(U)$ is called the *span* of U. Furthermore, if $\mathrm{sp}(U) = W$ then U is said *to span* W.

Remark 6. If U is a subset of a vector space V then $\mathrm{sp}(U)$ is a subspace of V.

Example 7. If $U = \{x^n : n \geq 0\} \subseteq C(-\infty, \infty)$ then $\mathrm{sp}(U) = \mathbb{R}[x]$.

Definition 7. A set U contained in a vector space V is called a *basis for* V if U is linear independent and $\mathrm{sp}(U) = V$.

Remark 7. It is possible to prove that every vector space has an basis. While we will be able to prove this for "finite dimensional vector spaces", the proof *must* for a general abstract vector space invoke the axiom of choice.

Proposition 1.4. *Let B be a basis for a vector space V over a field $\mathbb{F}$. If $\mathbf{b}_1, \ldots, \mathbf{b}_n \in B$, $c_1, \ldots, c_n, c_1', \ldots, c_n' \in \mathbb{F}$ are such that*

$$c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n = c_1'\mathbf{b}_1 + \cdots + c_n'\mathbf{b}_n$$

then $c_i = c_i'$.

Remark 8. Proposition 1.4 says that every vector has a unique representation as a linear sum of elements in a basis.

Lemma 1.5. *(Steinitz exchange lemma) Let V be a vector space and let S be a set with n elements such that $\mathrm{sp}(S) = V$. If T is a linearly independent subset of V with m elements then $m \leq n$ and there exists an $R \subseteq S$ containing $n - m$ elements such that $\mathrm{sp}(T \cup R) = V$.*

Theorem 1.6. *If B and B' are bases for a vector space V and B has $n \in \mathbb{N}$ elements then B' has n elements.*

Definition 8. A vector space V is called *finite dimensional* with dimension n if there exists a basis B for V with n elements. In this case we write $\dim V = n$. A space that is not finite dimensional is said to be *infinite dimensional*.

Example 8.
(a) $\mathbb{R}^n$ has dimension n.
(b) The vector space $M_{n\times m}(\mathbb{F})$ of $n \times m$ matrices with entries in $\mathbb{F}$ taken with the usual notion of matrix addition and scaling has dimension nm.
(c) The vector space $C[0,1]$ of continuous real valued functions on $[0,1]$ is infinite dimensional.

Proposition 1.7. *Let V be an n-dimensional vector space and let $S \subset V$ be a linearly independent set with k elements. The following statements are true.*
(a) $k \leq n$.
(b) If $k = n$ then S is a basis for V.
(c) There exists a basis B for V such that $S \subseteq B$.

Proposition 1.8. *Let V be a vector space and let U, W be finite dimensional subspaces of V. The following statements are true.*
(a) $\dim(U) \leq \dim(V)$ *with equality if and only if $U = V$.*

(b) $\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W)$.

Example 9. For $U = \mathrm{sp}\{1, x^2\}$ and $W = \mathrm{sp}\{1 - x^2, x\}$ subspaces of $C[0,1]$ we can see that $\dim(U \cap W) = 1$.

CHAPTER 2

Linear maps

1. Linear maps

Definition 9. Let V and W be vector spaces over a field $\mathbb{F}$. A function $L : V \to W$ is called a *linear transformation* or a *linear map* if for all $u, v \in V$ and $c \in \mathbb{F}$ we have

(a) $L(\mathbf{u} + \mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$

(b) $L(c\mathbf{v}) = cL(\mathbf{v})$.

When $V = W$, the linear transformation L is often called a *linear operator*.

PROPOSITION 2.1. *Let $L : V \to W$ be a map between vector spaces over a field $\mathbb{F}$. The following statements are true.*

(a) If L is linear then $L(\mathbf{0}) = \mathbf{0}$.

(b) L is linear if and only if $L(\mathbf{u} + c\mathbf{v}) = L(\mathbf{u}) + cL(\mathbf{v})$ for all $u, v \in V$ and $c \in \mathbb{F}$.

(c) If L is linear then $L(V)$ is a vector space.

Example 10. We now provide a list of examples of linear transformations.

(1) The map $I_V : V \to V$ defined by $\mathbf{v} \mapsto \mathbf{v}$ is a linear transformation. This linear transformation is called the *identity transformation.*

(2) For $A \in M_{n\times m}(\mathbb{F})$ the map defined by $\phi_A : \mathbf{x} \mapsto A\mathbf{x}$ defines a linear transformation

$$\phi_A : \mathbb{F}^m \to \mathbb{F}^n.$$

(3) The map $\int : f \mapsto \int f(x)\mathrm{dx}$ with the constant of integration taken to be zero defines a linear transformation

$$\int : C[a, b] \to C[a, b].$$

(4) The map $\frac{d}{dx} : f \mapsto \frac{d}{dx}f(x)$ defines a linear transformation

$$\tfrac{d}{dx} : C^\infty[a, b] \to C^\infty[a, b].$$

Remark 9. Let V and W be vector spaces over a field $\mathbb{F}$ and let $L : V \to W$ be a linear map. If $B = \{\mathbf{v}_i\}$ is a basis for V, then the linear map L is determined by $\{L(\mathbf{v}_i)\}$. Namely, for any $\mathbf{v} \in V$ there exists $c_{k_1}, \ldots, c_{k_n} \in \mathbb{F}$ so that $\mathbf{v} = c_{k_1}\mathbf{v}_{k_1} + \cdots + c_{k_n}\mathbf{v}_{k_n}$ and linearity tells us that

$$L(\mathbf{v}) = L(c_{k_1}\mathbf{v}_{k_1}+\cdots+c_{k_n}\mathbf{v}_{k_n}) = L(c_{k_1}\mathbf{v}_{k_1})+\cdots+L(c_{k_n}\mathbf{v}_{k_n}) = c_{k_1}L(\mathbf{v}_{k_1})+\cdots+c_{k_n}L(\mathbf{v}_{k_n}).$$

Definition 10. A linear transformation $L : V \to W$ between vector spaces is said to be *invertible* if there exists a map $L' : W \to V$ such that $L \circ L' = I_W$ and $L' \circ L = I_V$.

PROPOSITION 2.2. *Let $L : V \to W$ be a linear map. If L is invertible then the inverse map L' is a linear transformation and is unique.*

Remark 10. As the map L' from Definition 10 is unique, we call L' *the inverse* of L. The inverse of L is denoted L^{-1}.

Definition 11. Two vector spaces V and W are said to be *isomorphic* if there exists an invertible linear map $L : V \to W$. When V and W are isomorphic we denote this by writing $V \cong W$.

Example 11. The space $\mathbb{R}[x]_{\leq n}$ of polynomials of degree at most n is isomorphic to $\mathbb{R}^{n+1}$.

THEOREM 2.3. *If V is a vector space of dimension n over a field $\mathbb{F}$ then $V \cong \mathbb{F}^n$.*

Definition 12. Let $L : V \to W$ be a linear transformation between finite dimensional vector spaces. We define the *kernel of* L, denoted $\ker(L)$, and the *image of* L, denoted $\mathrm{im}(L)$, as follows:

- $\ker(L) := \{\mathbf{v} \in V : L(\mathbf{v}) = \mathbf{0}\}$.
- $\mathrm{im}(L) := \{\mathbf{w} \in W : \mathbf{w} = L(\mathbf{v}) \text{ for some } \mathbf{v} \in V\}$.

Example 12.

(a) The linear transformation $\frac{d}{dx} : \mathbb{R}[x] \to \mathbb{R}[x]$ defined by $p(x) \mapsto p'(x)$ has $\ker(\frac{d}{dx}) = \mathbb{R}$ and $\mathrm{im}(\frac{\mathrm{d}}{\mathrm{dx}}) = \mathbb{R}[\mathrm{x}]$.

(b) The linear transformation $P_i : \mathbb{R}^n \to \mathbb{R}^n$ defined by $c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n \mapsto c_i\mathbf{e}_i$ for $1 \leq i \leq n$ has $\ker(P_i) = \mathrm{sp}(\{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \backslash \{\mathbf{e}_i\})$ and $\mathrm{im}(P_i) = \mathrm{sp}\{\mathbf{e}_i\}$.

THEOREM 2.4. *If $L : V \to W$ is a linear transformation between vector spaces then* $\ker(L) \subset V$ *and* $\mathrm{im(L)} \subset \mathrm{W}$ *are subspaces.*

Definition 13. Let $L : V \to W$ be a linear transformation between vector spaces. If $\ker(L)$ and $\mathrm{im}(L)$ are finite dimensional then we call $\dim(\mathrm{im(L)})$ the *rank* of L and $\dim(\ker(L))$ the *nullity* of L.

THEOREM 2.5. *(Rank-Nullity Theorem) Let $L : V \to W$ be a linear transformation between vector spaces. If V is finite dimensional then*

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

Example 13. The map $L : \mathbb{R}[x]_{\leq n} \to \mathbb{R}[x]_{\leq n+1}$ defined by

$$f(x) \mapsto \int_0^x f(t)dt + f''(x)$$

has $\dim(\mathrm{im}(L)) = n + 1$.

PROPOSITION 2.6. *Let $L : V \to W$ be a linear transformation between vector spaces. The map L is one-to-one if and only if ker$(L) = \{\mathbf{0}\}$. Moreover, if $V = W$ and V is finite dimensional then L is one-to-one if and only if L is onto.*

Example 14. The map L from Example 13 is one-to-one but not onto.

2. Linear maps between finite dimensional vector spaces

PROPOSITION 2.7. *Let ϕ_A be as in Example 10. If $L : \mathbb{F}^m \to \mathbb{F}^n$ is a linear map then there exists a matrix $A \in M_{n\times m}(\mathbb{R})$ such that $L = \phi_A$.*

Remark 11. Theorem 2.3 tells us that every finite dimensional vector space is isomorphic to $\mathbb{F}^n$ and Proposition 2.7 tells us that every linear map between spaces of the form $\mathbb{F}^n$ is "like" a matrix. Linear maps are defined in terms of bases (see Remark 9). We will now show that all linear maps between finite dimensional vector spaces can be represented by matrices that depend only on the choice of bases[1].

Definition 14. Let $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be an ordered basis for a vector space V over a field $\mathbb{F}$ and let $c_1, \dots c_n \in \mathbb{F}$ be such that $\mathbf{v} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$. We define

$$[\mathbf{v}]_B := \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

We call $[\mathbf{v}]_B$ the *coordinate vector of* $\boldsymbol{v}$ *relative to the basis* B.

Example 15. With respect to the standard basis, the vector $3\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3$ is written as $\mathbf{v} = (3, 2, 1)^T \in \mathbb{R}^3$. The ordered set $B = \{(1, 1, 1), (0, 1, 1), (1, 0, 1)\}$ is an ordered basis for $V = \mathbb{R}^3$. With respect to the basis B, the coordinate vector of $\mathbf{v}$ is given by

$$[\mathbf{v}]_B := \begin{pmatrix} 4 \\ -2 \\ -1 \end{pmatrix}.$$

Example 16. $B = \{x^2 + 2x + 1, x + 1, x - 1\}$ is an ordered basis for $\mathbb{R}[x]_{\le 2}$. With respect to the basis B, the coordinate vector of $3x^2 + x + 6$ is given by

$$[3x^2 + x + 6]_B := \begin{pmatrix} 3 \\ -1 \\ -4 \end{pmatrix}.$$

PROPOSITION 2.8. *Let V and W be vector spaces with bases $B_V = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $B_W = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$, respectively. Let $L : V \to W$ be a linear transformation defined by*

$$\mathbf{v}_i \mapsto L(\mathbf{v}_i) = c_{i1}\mathbf{w}_1 + \cdots + c_{im}\mathbf{w}_m.$$

We define the matrix $[L]_{B_V}^{B_W}$ by

$$[L]_{B_V}^{B_W} := \begin{pmatrix} | & & | \\ [L(\mathbf{v}_1)]_{B_W} & \dots & [L(\mathbf{v}_n)]_{B_W} \\ | & & | \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{n1} \\ \vdots & \ddots & \vdots \\ c_{1m} & \dots & c_{nm} \end{pmatrix}.$$

[1] the choice of bases on $\mathbb{F}^m$ and $\mathbb{F}^n$ in Example 10/Proposition 2.7 are the standard bases

For $\mathbf{v} \in V$ *we have the following identity:*

$$[L]_{B_V}^{B_W}[\mathbf{v}]_{B_V} = [L(\mathbf{v})]_{B_W}.$$

Remark 12. Multiplication by the matrix $[L]_{B_V}^{B_W}$ takes a vector $\mathbf{v} \in V$, written as a coordinate vector with respect to the basis B_V, and returns $L(v) \in W$ as a coordinate vector with respect to the basis B_W.

Definition 15. Let V and W be vector spaces with bases $B_V = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $B_W = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ respectively. The matrix $[L]_{B_V}^{B_W}$ defined in Proposition 2.9 is called the *matrix representation of* L *relative to the bases* B_V *and* B_W.

Example 17. Let $\phi_A : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map associated to the matrix $A \in M_{2\times 2}$ given by

$$A := \begin{pmatrix} 2 & 3 \\ 3 & -6 \end{pmatrix}.$$

$B = \{(1,3)^T, (-3,1)^T\}$ is a basis for $\mathbb{R}^2$ and

$$[L]_B^B = \begin{pmatrix} 3 & 0 \\ 0 & -7 \end{pmatrix}.$$

Remark 13. The vectors in B from Example 17 are eigenvectors of A and the diagonal entries of $[L]_B^B$ are the corresponding eigenvalues.

Example 18. Let $\frac{\mathrm{d}}{\mathrm{dx}} : \mathbb{R}[x]_{\leq 3} \to \mathbb{R}[x]_{\leq 2}$ be defined by $p(x) \mapsto p'(x)$. Taking $B = \{x^3, x^2, x, 1\}$ as a basis for $\mathbb{R}[x]_{\leq 3}$ and $C = \{1, 1+x, 1+x^2\}$ we find that the matrix representation of $\frac{\mathrm{d}}{\mathrm{dx}}$ relative to B and C is

$$\left[\frac{\mathrm{d}}{\mathrm{dx}}\right]_B^C = \begin{pmatrix} | & | & | & | \\ [3x^2]_C & [2x]_C & [1]_C & [0]_C \\ | & | & | & | \end{pmatrix} = \begin{pmatrix} -3 & -2 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{pmatrix}$$

Everything that we would expect for the matrix/coordinate representations of linear maps/vectors is true.

PROPOSITION 2.9. *Let $I_V : V \to V$ be the identity map between vector spaces, and let $B = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $B' = \{\boldsymbol{v}'_1, \dots, \boldsymbol{v}'_n\}$ be basis for V. For every $\mathbf{v} \in V$ we have*

$$[I_V]_B^{B'}[\mathbf{v}]_B = [\mathbf{v}]_{B'}.$$

Definition 16. The matrix $[I_V]_B^{B'}$ in Proposition 2.9 is called the *change of basis matrix.*

Example 19. If B is a basis of an n dimensional vector space V, then the matrix $[I_V]_B^B$ is the standard $n \times n$ identity matrix.

Example 20. Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $C = \{\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3\}$ be basis for a vector space V. The change of basis matrix is given by

$$[I_V]_B^C = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

PROPOSITION 2.10. *Let $L : V \to W$ be linear transformation between finite dimensional vector spaces and let B be a basis for V, and C be a basis for W. The following statements are true:*
(a) L is invertible if and only if $[L]_B^C$ is invertible
(b) If L is invertible then $[L^{-1}]_C^B = ([L]_B^C)^{-1}$

Example 21. The identity operator $I_V : V \to V$ is invertible. Taking B and C from Example 20 we can easily verify that $[I_V]_B^C$ is invertible and that $[I_V]_C^B = ([I_V]_B^C)^{-1}$.

PROPOSITION 2.11. *Let $L : V \to W$ be a linear transformation between finite dimensional vector spaces and let B_1, B_2 be bases for V, and C_1, C_2 be bases for W. The following statement is true:*

$$[L]_{B'}^{C'} = [I_W]_C^{C'}[L]_B^C[I_V]_{B'}^B$$

Example 22. Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $C = \{\mathbf{w}_1, \mathbf{w}_2\}$ be bases for vector spaces V and W respectively, and let $L : V \to W$ be a linear transformation defined by

$$\mathbf{v}_1 \mapsto \mathbf{w}_1 + 2\mathbf{w}_2, \qquad \mathbf{v}_2 \mapsto 2\mathbf{w}_1 + \mathbf{w}_2, \qquad \mathbf{v}_3 \mapsto \mathbf{w}_1 + \mathbf{w}_2.$$

$B' = \{\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 + \mathbf{v}_3, \mathbf{v}_3 + \mathbf{v}_1\}$ is a basis for V and $C' = \{\mathbf{w}_1 + \mathbf{w}_2, \mathbf{w}_1 - \mathbf{w}_2\}$ is a basis for W. We find

$$[L]_{B'}^{C'} = [I_W]_C^{C'}[L]_B^C[I_V]_{B'}^B = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{3}{2} & \frac{5}{2} & \frac{5}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

Definition 17. Let $L : V \to V$ be a linear operator on a vector space V over a field $\mathbb{F}$. A vector $\mathbf{v} \in V \backslash \{\mathbf{0}\}$ is called an *eigenvector* if there exists a $\lambda \in \mathbb{F}$ such that $L(\mathbf{v}) = \lambda \mathbf{v}$. The scalar λ is called the *eigenvalue* corresponding to the eigenvector $\mathbf{v}$.

Example 23. Let $\phi_A : \mathbb{R}^n \to \mathbb{R}^n$ be the linear operator associated to a matrix $A \in M_{n\times n}(\mathbb{F})$. The eigenvectors and eigenvalues of A correspond to the eigenvectors and eigenvalues of ϕ_A.

Example 24. Let $L : \mathbb{R}[x]_{\leq 5} \to \mathbb{R}[x]_{\leq 5}$ be the linear operator defined by $p(x) \mapsto p(-x)$. The elements of $U = \mathrm{sp}\{1, x^2, x^4\}$ are eigenvectors of L with eigenvalue 1 and the elements of $U' = \mathrm{sp}\{x, x^3, x^5\}$ are eigenvectors of L with eigenvalue -1.

Using Propositions 2.10 and 2.11 we will now see that we can define determinant, trace, characteristic polynomial, and compute eigenvectors and eigenvalues of a linear operator on a finite dimensional vector space.

PROPOSITION 2.12. *Let $L : V \to V$ be a linear operator on a finite dimensional vector space V and let B, B' be bases for V. The following statements are true*
(a) $\det([L]_B^B) = \det([L]_{B'}^{B'})$
(b) $\mathrm{trace}([L]_B^B) = \mathrm{trace}([L]_{B'}^{B'})$
(c) $\det([L]_B^B - \lambda[I_V]_B^B) = \det([L]_{B'}^{B'} - \lambda[I_V]_{B'}^{B'})$

Proposition 2.12 means that the following definitions for a linear operator on a finite dimensional vector space are independent of the choice of basis.

Definition 18. Let $L : V \to V$ be a linear operator on a finite dimensional vector space V and let B be a basis for V. We define

- The *determinant* of L, denoted $\det(L)$, by $\det([L]_B^B)$.
- The *characteristic polynomial* of L is defined by

$$\chi_L(\lambda) := \det([L]_B^B - \lambda [I_V]_B^B).$$

The fact that the characteristic polynomial of a linear operator is well defined means that we have a method for computing eigenvalues and eigenvectors of a linear operator.

Example 25. Let $L : R[x]_{\leq 2} \to \mathbb{R}[x]_{\leq 2}$ be the linear map defined by

$$1 \mapsto 1, \qquad x \mapsto 1 + 2x, \qquad x^2 \mapsto 1 - x^2.$$

Using $B = \{1, x, x^2\}$ as a basis for $R[x]_{\leq 2}$ we find

$$[L]_B^B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

It is now easy to check that $\det(L) = -2$, $\mathrm{trace}(L) = 2$. The characteristic polynomial of L is given by $\chi_L(\lambda) = (1-\lambda)(2-\lambda)(-1-\lambda)$ and $1, 1+x, -1+2x^2$ are eigenvectors corresponding to the eigenvalues $1, 2, -1$ respectively.

CHAPTER 3

Inner product spaces

1. Basics

In $\mathbb{R}^n$ we have a natural notion of an angle between two vectors. By adding some addition structure on an abstract vector space we will be able to abstract the notion of an "angle" between two elements of a general vector space.

Definition 19. Let V be a vector space over a field $\mathbb{F} \subseteq \mathbb{C}$. An *inner product* is a map $\langle , \rangle : V \times V \to \mathbb{F}$ given by $(\mathbf{u}, \mathbf{v}) \mapsto \langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{F}$ satisfying the following properties:

IP1: $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.
IP2: $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$ for all $\mathbf{u}, \mathbf{v} \in V$ and $c \in \mathbb{F}$.
IP3: $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$ for all $\mathbf{u}, \mathbf{v} \in V$.
IP4: We have $\langle \mathbf{v}, \mathbf{v} \rangle \in \mathbb{R}$ with $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ for all $\mathbf{v} \in V \backslash \{\mathbf{0}\}$.

A vector space over $\mathbb{F}$ together with a map $\langle , \rangle : V \times V \to \mathbb{F}$ satisfying **IP1**-**IP4** is called an *inner product space.*

Definition 20. When we take $\mathbb{F} = \mathbb{R}$ in the definition of an inner product space, Axioms **IP2** and **IP3** simplify to

IP2′: $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$ for all $\mathbf{u}, \mathbf{v} \in V$ and $c \in \mathbb{R}$.
IP3′: $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ for all $\mathbf{u}, \mathbf{v} \in V$.

A vector space over $\mathbb{R}$ with an inner product $\langle , \rangle : V \times V \to \mathbb{R}$ satisfying **IP1**, **IP2′**, **IP3′**, **IP4** is called a *real inner product space.*

Remark 14. Throughout this section we will almost exclusively work with real inner product spaces. However, all of the arguments will work (with minor modifications) for vector spaces over $\mathbb{C}$.

Example 26. $V = \mathbb{R}^n$ with $\langle (x_1, \ldots, x_n)^T, (y_1, \ldots, y_n)^T \rangle := \sum_{i=1}^n x_i y_i$ is a real inner product space.

Example 27. $V = C[a,b]$ the space of continuous real valued functions over the interval $[a,b]$ with

$$\langle f(x), g(x)\rangle := \int_a^b f(x)g(x)\mathrm{dx}$$

is a real inner product space.

Example 28. All subspaces of the inner product space $C[a,b]$ are inner product spaces with respect to the inner product defined in Example 27. In particular, $V = C^p$, the space of p-periodic[1] continuous functions $f : \mathbb{R} \to \mathbb{R}$, with

$$\langle f(x), g(x)\rangle := \int_{-\frac{p}{2}}^{\frac{p}{2}} f(x)g(x)\mathrm{dx}$$

is a real inner product space.

PROPOSITION 3.1. *Let V be an inner product space over $\mathbb{R}$. For $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $c \in \mathbb{R}$ the following statements are true:*

(1) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle$.
(2) $\langle \mathbf{u}, c\mathbf{v}\rangle = c\langle \mathbf{u}, \mathbf{v}\rangle$.
(3) $\langle \mathbf{u}, \mathbf{u}\rangle = 0$ *if and only if* $\mathbf{u} = \mathbf{0}$.
(4) if $\langle \mathbf{x}, \mathbf{u}\rangle = \langle \mathbf{x}, \mathbf{v}\rangle$ *for all* $\mathbf{x} \in V$ *then* $\mathbf{u} = \mathbf{v}$.

Definition 21. Let V be an inner product space. We define

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

$\|\mathbf{v}\|$ is called the *length* or *norm* of the vector $\mathbf{v} \in V$.

Example 29.
(a) For $\mathbf{v} = (1, 2, -7) \in \mathbb{R}^3$ we have $\|\mathbf{v}\| = \sqrt{1^2 + 2^2 + (-7)^2} = 3\sqrt{6}$.
(b) For $x^3 \in C[0,1]$ we have

$$\|x^3\| = \sqrt{\int_0^1 (x^3)^2 dx} = \tfrac{1}{\sqrt{7}}.$$

PROPOSITION 3.2. *Let V be an inner product space. For $\mathbf{u}, \mathbf{v} \in V$ and $c \in \mathbb{R}$ the following statements hold:*
(a) $\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|$.

[1] a function $f : \mathbb{R} \to \mathbb{R}$ is called p-periodic if $f(x + p) = f(x)$ for all $x \in \mathbb{R}$

(b) $\|\mathbf{v}\| \geq 0$ *and* $\|\mathbf{v}\| = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$.

THEOREM 3.3. *(Cauchy-Schwarz inequality) Let* V *be an inner product space. For* $\mathbf{u}, \mathbf{v} \in V$ *the following inequality holds:*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\| \, .$$

Furthermore, $|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\| \, \|\mathbf{v}\|$ *if and only if* $\mathbf{u}$ *is a scalar multiple of* $\mathbf{v}$.

Example 30. For $f(x), g(x) \in C[a, b]$ we have

$$\left| \int_a^b f(x)g(x)dx \right|^2 \leq \left(\int_a^b (f(x))^2 dx \right) \left(\int_a^b (g(x))^2 dx \right).$$

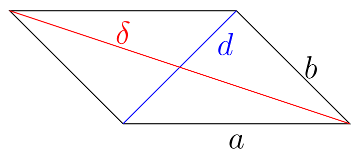
THEOREM 3.4. *(triangle inequality) Let* V *be an inner product space. For* $\mathbf{u}, \mathbf{v} \in V$ *the following inequality holds:*

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\| \, .$$

PROPOSITION 3.5. *(parallelogram law) Let* V *be an inner product space. For* $\mathbf{u}, \mathbf{v} \in V$ *the following identity holds:*

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

Example 31. Applying the parallelogram law to $\mathbb{R}^2$ we find that the sum of the squares of the edges of a parallelogram is equal to the sum of the squares of the diagonals. i.e. $a^2 + b^2 + a^2 + b^2 = d^2 + \delta^2$ in the diagram below.



Definition 22. Two vectors $\mathbf{u}, \mathbf{v}$ in an inner product space V are said to be *orthogonal* or *perpendicular* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

THEOREM 3.6. *(Pythagoras' theorem)* $\mathbf{u}, \mathbf{v}$ *are orthogonal vectors in a real inner product space if and only if*

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \, .$$

2. Linear maps

Given an inner product on a vector space we are able to impose additional structure on linear maps. We begin with a theorem.

THEOREM 3.7. *Let V be an inner product space over a field $\mathbb{F}$ and let $L : V \to V$ be a linear transformation. If there exists a map $L^* : V \to V$ with $\langle L(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, L^*(\mathbf{v})\rangle$ for all $\mathbf{u}, \mathbf{v} \in V$, then L^* is linear and unique. Furthermore, if V is finite dimensional then L^* exists.*

Definition 23. If the linear transformation L^* from Theorem 3.7 exists then it is called the *adjoint* of L.

Example 32. If $V = \mathbb{R}^n$ and $\phi_A : \mathbb{R}^n \to \mathbb{R}^n$ is the linear map associated to a matrix $A \in M_{n\times n}(\mathbb{R})$ then $(\phi_A)^* = \phi_{A^T}$.

Remark 15. As the adjoint of a linear operator is unique, Example 32 show that $B = A^T$ is the unique $n \times n$ matrix satisfying $\mathbf{x}^T A\mathbf{y} = \mathbf{y}^T B\mathbf{x}$.

PROPOSITION 3.8. *Let V be an inner product space and let L, N be linear operators on V. If L^* and T^* exist then the following hold:*
(a) $(L + T)^* = L^* + T^*$.
(b) $(cL)^* = cL^*$ *for* $c \in \mathbb{R}$ *when V is a real inner product vector space.*
(c) $(LT)^* = T^*L^*$.
(d) $(L^*)^* = L$.

Remark 16. Proposition 3.8 holds when L/T are $n \times n$ matrices and "L^*/T^*" are read as "L^T/T^T" (see Example 32).

Definition 24. A linear operator $L : V \to V$ on an inner product space V is said to be *self-adjoint* if $L = L^*$.

Example 33. If $L : V \to V$ is a linear operator with an adjoint L^* then $L + L^*$ is self-adjoint operator on V.

PROPOSITION 3.9. *All eigenvalues of a self-adjoint linear operator are real.*

Definition 25. Let $A \in M_{n\times n}(\mathbb{C})$ whose ij^{th} entry is denoted a_{ij}. We define $A^\dagger$ to be the matrix whose ij^{th} entry of is $\overline{a_{ji}}$.

Example 34. Let $\phi_A : \mathbb{C}^n \to \mathbb{C}^n$ be the linear operator associated to a matrix $A \in M_{n\times n}(\mathbb{C})$ with $A = A^\dagger$. It turns out that $(\phi_A)^* = \phi_{A^\dagger}$. So, Proposition 3.9 tells us that the eigenvalues of the matrix $A = A^\dagger$ are all real.

Lemma 3.10. *If W is a subspace of an inner product space V and $L : V \to V$ is a self-adjoint operator with $L(W) \subseteq W$ then $L(W^\perp) \subseteq W^\perp$.*

Theorem 3.11. *If $L : V \to V$ is a self-adjoint operator on an n-dimensional vector space then there exists an orthonormal basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ with each $\mathbf{v}_i$ an eigenvector of L.*

Remark 17. An immediate consequence of Theorem 3.11 is that if $A \in M_{n\times n}(\mathbb{C})$ is a such that $A = A^\dagger$ then there exists an orthonormal basis for $\mathbb{C}^n$ whose elements are eigenvectors of A.

CHAPTER 4

Gram-Schmidt and Fourier theory

1. Gram-Schmidt

In $\mathbb{R}^n$ the angle between two vectors $\mathbf{u}, \mathbf{v}$ is given by $\arccos\left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}\right)$. So, two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are perpendicular if and only if $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}\cdot\mathbf{v} = 0$. This motivates the following definition.

Definition 26. Let V be a real inner product space. We make the following definitions:

- A set $S \subset V$ is said to be *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for every pair of distinct $\mathbf{u}, \mathbf{v} \in S$.
- A vector $\mathbf{v} \in V$ is called a *unit* vector if $\|\mathbf{v}\| = 1$.
- A set $S \subset V$ is said to be *orthonormal* if S is orthogonal and each element of S is a unit.
- A set $B \subset V$ is called an *orthonormal/orthogonal basis* for V if B is a basis for V and B is orthonormal/orthogonal.

Example 35. $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is an orthonormal basis for $\mathbb{R}^n$.

Example 36. The set $\{1, \sin(x), \cos(x), \ldots, \sin(nx), \cos(nx)\}$ is an orthogonal set in $C^{2\pi}$.

THEOREM 4.1. *(Gram-Schmidt orthogonalisation) Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a linearly independent set of vectors in a real inner product space V and define*

$$\mathbf{u}_1 := \mathbf{v}_1, \qquad \textit{and} \qquad \mathbf{u}_i := \mathbf{v}_i - \sum_{j=1}^{i-1} \frac{\langle \mathbf{v}_i, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \mathbf{u}_j \qquad \textit{for } i \geq 2.$$

The set $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is orthogonal and $sp(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}) = sp(\{\mathbf{u}_1, \ldots, \mathbf{u}_k\})$.

Remark 18. By normalising the vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ in Theorem 4.1 the Gram-Schmidt method can be used to obtain an orthonormal set $\left\{\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \ldots, \frac{\mathbf{u}_k}{\|\mathbf{u}_n\|}\right\}$ with

$$\operatorname{sp}(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}) = \operatorname{sp}\left(\left\{\frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \ldots, \frac{\mathbf{u}_k}{\|\mathbf{u}_n\|}\right\}\right).$$

Example 37. We know that $\{1, x, x^2\} \subseteq C[0,1]$ is a linearly independent set. We can now use the Gram-Schmidt method to find $\{1, x - \frac{1}{2}, x^2 - x - \frac{1}{6}\}$ as an orthogonal basis for $\operatorname{sp}(\{1, x, x^2\})$ and $\{1, 12x - 6, 6\sqrt{5}x^2 - 6\sqrt{5}x - \sqrt{5}\}$ as an orthonormal basis for $\operatorname{sp}(\{1, x, x^2\})$.

Definition 27. Let S be a subset of an inner product space V. We define

$$S^\perp := \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \in S\}.$$

Remark 19. $S^\perp$ is read “S perp”.

Example 38. For linearly independent $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ and $U = \operatorname{sp}(\{\mathbf{x}, \mathbf{y}\})$ we have $U^\perp = \operatorname{sp}\{\mathbf{x} \times \mathbf{y}\}$.

PROPOSITION 4.2. *If S is a subset of an inner product space V then the following statements are true:*
(a) $S^\perp$ is a subspace of V.
(b) $\operatorname{sp}(S) \cap S^\perp = \{\mathbf{0}\}$.
(c) If S is an orthogonal set then S is a linearly independent set.
(d) For all $\mathbf{v} \in V$, if $\operatorname{sp}(S)$ is finite dimensional then there exists a unique $\mathbf{x} \in \operatorname{sp}(S)$ and $\mathbf{y} \in S^\perp$ such that $\mathbf{v} = \mathbf{x} + \mathbf{y}$.

Definition 28. Let U be a subspace of a real inner product space V. For $\mathbf{v} \in V$ we let $\mathbf{u} \in U$ and $\mathbf{w} \in U^\perp$ be the unique vectors with $\mathbf{v} = \mathbf{u} + \mathbf{w}$. The vector $\mathbf{u}$ is called *the projection of* $\mathbf{v}$ *onto* U, and is denoted $\operatorname{Proj}_U(\mathbf{v})$.

THEOREM 4.3. *Let U be a finite dimensional subspace of a real inner product space V and let $\mathbf{v} \in V$. The following identity is true*

$$\|\mathbf{v} - \operatorname{Proj}_U(\mathbf{v})\| = \inf\{\|\mathbf{v} - \mathbf{w}\| : \mathbf{w} \in U\}.$$

Moreover, if $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ *is an orthogonal basis for* U *then*

$$\mathrm{Proj}_U(\mathbf{v}) = \sum_{i=1}^{k} \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$

Remark 20. If $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthonormal basis for U then

$$\mathrm{Proj}_U(\mathbf{v}) = \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i.$$

Example 39. Let $U = \mathrm{sp}\{1, x\}$ be a subspace of $C[0,1]$. We know that $\mathrm{Proj}_U(x^4) = a + bx$ corresponds to the minimum value in the set

$$\left\{ \int_0^1 (x^4 - a - bx)^2 \mathrm{dx} : a, b \in \mathbb{R} \right\}.$$

2. Fourier theory

This section will concentrate on the real inner product space $C^{2\pi}$ of continuous 2π periodic functions $f : \mathbb{R} \to \mathbb{R}$ with the inner product defined by

$$\langle f(x), g(x) \rangle := \int_{-\pi}^{\pi} f(x) g(x) \mathrm{dx}.$$

Definition 29. We define $T_n := \mathrm{sp}\{1, \sin(x), \cos(x), \dots, \sin(nx), \cos(nx)\}$.

Example 40. The set T_n is orthogonal in $C^{2\pi}$. For $f(x) \in C^{2\pi}$ we have

$$\begin{aligned}
\mathrm{Proj}_{T_n}(f(x)) &= \sum_{k=0}^{n} \frac{\langle f(x), \cos(kx) \rangle}{\langle \cos(kx), \cos(kx) \rangle} \cos(kx) + \sum_{k=1}^{n} \frac{\langle f(x), \sin(kx) \rangle}{\langle \sin(kx), \sin(kx) \rangle} \sin(kx). \\
&= \frac{a_0}{2} + \sum_{k=1}^{n} (a_n \cos(kx) + b_n \sin(kx))
\end{aligned}$$

$$\text{where} \qquad a_k := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) \mathrm{dx} \quad \text{and} \quad b_k := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) \mathrm{dx}$$

Remark 21. Example 40 tells us that the Fourier polynomial of order n of a continuous 2π-periodic function $f(x)$ is $\mathrm{Proj}_{T_n}(f(x))$.

We can state Fourier's theorem and interpret it as a statement about elements in an inner product space.

THEOREM 4.4. *(Convergence of Fourier series) For any* $f(x) \in C^{2\pi}$ *we have* $\|f(x) - \mathrm{Proj}_{T_n}(f(x))\| \to 0$ *as* $n \to \infty$.

Remark 22. Theorem 4.4 says that for sufficiently large n the distance from a function $f(x) \in C^{2\pi}$ to a function in T_n is as close as you like.

Remark 23. $C^{2\pi}$ can be viewed as a subspace of the inner product space of "Lebesque square integrable" functions over $[-\pi, \pi]$, denoted "$L^2[-\pi, \pi]$", and Theorem 4.4 holds in this more general setting.

Remark 24. The proof of Theorem 4.4 is not examinable. However, full details are available in a separate document.