SPRING 2023 MATH 5425 SUPPLEMENTARY NOTES

Contents

1. Vector Spaces and Subspaces

Roughly speaking, a real vector space is a set of vectors together with rules for vector addition and multiplication by real numbers. These operations must produce vectors in the space and must satisfy some conditions.

More precisely, consider a set V endowed with two operations:

- *Addition*: given two vectors v_1 and v_2 in V, it associates a new vector in V, denoted by v_1+v_2; i.e.,

$$\begin{aligned} + : V \times V &\to V \\ (v_1, v_2) &\mapsto v_1 + v_2. \end{aligned} \tag{1.1}$$

- *Scalar multiplication*: given a vector v in V and a real number c, it associates a new vector in V, denoted by $c \cdot v$ (or simply cv); i.e.,

$$\begin{aligned} \cdot : \mathbb{R} \times V &\to V \\ (c, v_2) &\mapsto c \cdot v. \end{aligned} \tag{1.2}$$

Definition 1.1. *A triple $(V, +, \cdot)$ is a (real) vector space if:*

(i) *for any v_1, v_2, $v_3 \in V$, $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$ (associative law);*
(ii) *there exists a zero vector $0 \in V$, such that $v + 0 = 0 + v = v$ for any $v \in V$;*
(iii) *for any $v \in V$, there exists $-v \in V$ satisfying $-v + v = v + (-v) = 0$;*
(iv) *for any $v_1, v_2 \in V$, $v_1 + v_2 = v_2 + v_1$ (commutative law);*
(v) *for any $v \in V$ and $c, d \in \mathbb{R}$, $(c + d) \cdot v = c \cdot v + d \cdot v$ (distributive law I);*
(vi) *for any $v_1, v_2 \in V$ and $c \in \mathbb{R}$, $c \cdot (v_1 + v_2) = c \cdot v_1 + c \cdot v_2$ (distributive law II);*
(vii) *for any $v \in V$ and $c, d \in \mathbb{R}$, $(cd) \cdot v = c \cdot (d \cdot v)$;*
(viii) *for any $v \in V$, $1 \cdot v = v$.*

The elements in V are called *vectors* and the elements in $\mathbb{R}$ *scalars*. When the scalar multiplication is defined on $\mathbb{C} \times V$, then the corresponding vector space is called a **complex vector space**. Unless specified, vector spaces throughout this note are assumed to be real vector spaces.

Example 1.1. *The following are some examples of vector spaces:*

(1) *$\mathbb{R}$ itself is a vector space;*

(2) *for any positive integer* n, $\mathbb{R}^n$ *(i.e., the set of real* n*-tuples) is a vector space;*
(3) *in general, if* V *is a vector space,* V^n *(i.e., the set of* n*-tuples whose components are elements of* V*) is still a vector space;*
(4) *the space of functions* $f : [0, 1] \to \mathbb{R}$ *is a vector space.*

Another important concept is the concept of *subspace*. A subspace of a vector space V is a subset of V that is closed under addition and scalar multiplication (roughly speaking, the results of the operations remain in this subset). These operations follow the rule of the host space, keeping us inside the subspace.

Definition 1.2. *Let* $(V, +, \cdot)$ *be a vector space and let* $W \subset V$ *a nonempty subset of* V*.* W *is said to be a vector subspace of* V*, if it is a vector space with the induced operations; i.e.* $(W, +, \cdot)$ *still satisfies* (i)–(viii) *in Definition 1.1, where* $+$ *and* $\cdot$ *are the same operations of* V*, restricted to* W*.*

A useful characterization of subspaces is the following:

Proposition 1.1. *Let* W *be a nonempty subset of a vector space* V*.* W *is a vector subspace if and only if the following two conditions are satisfied:*
(i) $w_1 + w_2 \in W$ *for any* $w_1, w_2 \in W$*;*
(ii) $c \cdot w \in W$ *for any* $c \in \mathbb{R}$ *and* $w \in W$*.*

Example 1.2. *The following are some examples of vector subspaces for a vector space* $(V, +, \cdot)$*.*
(1) *The subset containing only the zero vector,* $Z := \{0\}$ *and the whole space* V *are trivial subspaces (*Z *is the smallest possible subspace and* V *the biggest one).*
(2) *For any* $v_0 \in V$*, the set*
$$\langle v_0 \rangle := \{c \cdot v_0, c \in \mathbb{R}\}$$
is a vector subspace. This subspace is called subspace generated by v_0*. Observe that* $Z = \langle 0 \rangle$*.*
More generally, if $v_1, \cdots, v_k$ *are* k *fixed vectors in* V*, one can consider the set*
$$\langle v_1, \cdots, v_k \rangle := \{c_1 \cdot v_1 + \cdots + c_k \cdot v_k, \quad \forall\, c_1, \cdots, c_k \in \mathbb{R}\}.$$
This is a vector subspace and is called subspace generated by $v_1, \cdots, v_k$*.*
(3) *If* W_1 *and* W_2 *are two vector subspaces, then the intersection space* $W_1 \cap W_2$ *is still a subspace (the intersection is the subset containing all the elements that are both in* W_1 *and* W_2*).*
(4) *If* W_1 *and* W_2 *are two vector subspaces, then the sum space* $W_1 + W_2 := \{w_1 + w_2, w_1 \in W_1, w_2 \in W_2\}$*, is a vector subspace.*
Observe that $W_1 \cup W_2 \subset W_1 + W_2$*. In fact, in general the union might not be a subspace. One can show that* $W_1 + W_2$ *is the smallest subspace containing the union* $W_1 \cup W_2$*.*

Proposition 1.2. *Let* $(V, +, \cdot)$ *be a vector space and* W_1, W_2 *two subspaces. The following two conditions are equivalent:*
(i) $W_1 \cap W_2 = 0$*;*
(ii) *for any* $w \in W_1 + W_2$*, there exists a unique couple* $(w_1, w_2) \in W_1 \times W_2$*, such that* $w = w_1 + w_2$ *(i.e., any vector in the sum space can be written uniquely as the sum of vectors in* W_1 *and* W_2*).*

Definition 1.3. *Let* W_1, W_2 *be two vector subspaces of a vector space* V*.*
- *If* $W_1 \cap W_2 = 0$*, then the sum subspace* $W_1 + W_2$ *is called direct sum and it is denoted* $W_1 \oplus W_2$*.*
- W_1 *and* W_2 *are said* **supplementary subspaces** *if* $W_1 + W_2 = V$ *and* $W_1 \cap W_2 = 0$*.*

From Proposition 1.2 above, it follows immediately that to say that W_1 *and* W_2 *are supplementary subspaces* is equivalent to say that *every vector in* V *can be decomposed uniquely as the sum of a vector in* W_1 *and one in* W_2*.*

2. Linear Independence, Set of Generators and Basis

Definition 2.1. *Let $(V, +, \cdot)$ be a vector space, $v_1, \cdots, v_k \in V$ and $c_1, \cdots, c_k \in \mathbb{R}$. The vector*

$$c_1 \cdot v_1 + \cdots + c_k \cdot v_k$$

is called linear combination of $v_1, \cdots, v_k$ with coefficients $c_1, \cdots, c_k$.

Remark 2.1. *Observe that for any k vectors $v_1, \cdots, v_k \in V$, obviously*

$$0 \cdot v_1 + \cdots + 0 \cdot v_k = 0,$$

i.e., the zero vector is linear combination of the vectors $v_1, \cdots, v_k$ (with coefficients all equal to zero). The linear combination $0 \cdot v_1 + \cdots + 0 \cdot v_k$ is called trivial linear combination of $v_1, \cdots, v_k$.

A natural question is: given any k vectors $v_1, \cdots, v_k \in V$, is it possible to obtain the zero vector as a non-trivial linear combination of them (i.e., with at least one of the coefficients different from zero)?

Example 2.1. *Consider the following examples.*

(1) *In $\mathbb{R}^2$, the two vectors $v_1 = (1, 2)$ and $v_2 = (-2, -4)$ admits the following non-trivial linear combination*

$$2 \cdot v_1 + 1 \cdot v_2 = 0.$$

(2) *In $\mathbb{R}^2$, consider the vectors $v_1 = (1, 2)$ and $v_2 = (0, 1)$. Suppose that*

$$0 = c_1 \cdot v_1 + c_2 \cdot v_2 = c_1 \cdot (1, 2) + c_2 \cdot (0, 1) = (c_1, 2c_1 + c_2).$$

It follows that $c_1 = 0$ and $2c_1 + c_2 = 0$, and consequently $c_1 = c_2 = 0$. Hence, v_1 and v_2 do not admit a non-trivial combination of 0.

These examples justify the following definition.

Definition 2.2. *Let V be a vector space and let $v_1, \cdots, v_k \in V$. These vectors are said to be linearly independent if the only linear combination that gives the zero vector is the trivial one:*

$$c_1 \cdot v_1 + \cdots + c_k \cdot v_k = 0 \qquad \Longrightarrow \qquad c_1 = \cdots = c_k = 0.$$

Otherwise, they are said to be linearly dependent.

Remark 2.2.

(i) *A single vector $v \in V$ is linearly dependent if and only if $v = 0$.*
(ii) *Given k linearly independent vectors, any their subset still consists of linearly independent vectors.*
(iii) *From a geometrical point of view, we can think of the linear dependence in the following way:*
   - *two vectors v_1 and v_2 are linearly dependent if and only if they are parallel;*
   - *three vectors v_1, v_2 and v_3 are linearly dependent if and only if they are coplanar (i.e., contained in a same plane).*

Proposition 2.1. *Let $v_1, \cdots, v_k \in V$, with $k \geq 2$. They are linearly dependent if and only if at least one of them can be written as a linear combination of the remainings.*

Proposition 2.2. *Let $v_1, \cdots, v_k \in V$. They are linearly independent if and only if the following condition holds:*

$$\text{if} \quad \sum_{i=1}^{k} c_i \cdot v_i = \sum_{i=1}^{k} d_i \cdot v_i \qquad \Longrightarrow \qquad c_1 = d_1, \quad \cdots, \quad c_k = d_k. \tag{2.1}$$

Let us now introduce the concept of set of generators, or what means for a set of vectors to *span* a space. We have defined in Example 1.2 above the space generated by k vectors. If $v_1, \cdots, v_k$ are k fixed vectors in V, one can consider the set

$$\langle v_1, \cdots, v_k \rangle := \{c_1 \cdot v_1 + \cdots + c_k \cdot v_k, \quad \forall\, c_1, \cdots, c_k \in \mathbb{R}\}.$$

This is a vector subspace and is called **subspace generated by** $v_1, \cdots, v_k$. Observe that $\langle v_1, \cdots, v_k \rangle$ is the smallest vector subspace containing $v_1, \cdots, v_k$.

Definition 2.3. *Let V be a vector space and $v_1, \cdots, v_k \in V$. If*

$$\langle v_1, \cdots, v_k \rangle = V,$$

we say that $\{v_1, \cdots, v_k\}$ is a set of generators of V (or a spanning set of V).

More generally, let W be a subspace of V and let $w_1, \cdots, w_s \in W$. If

$$\langle w_1, \cdots, w_s \rangle = W,$$

we say that $\{w_1, \cdots, w_s\}$ is a set of generators of W (or a spanning set of W).

We have introduced all the ingredients that we need to define a basis of a vector space. In few words, a basis of a vector space is a set of generators, made of linearly independent vectors. More formally:

Definition 2.4. *Let V be a vector space and let $v_1, \cdots, v_n \in V$. We say that $\{v_1, \cdots, v_n\}$ is a* **basis** *of V, if:*

(i) *$v_1, \cdots, v_n$ are linearly independent;*
(ii) *$\langle v_1, \cdots, v_n \rangle = V$.*

Why should one interested in considering a basis on a vector space? A basis allows us to associate to each vector its coordinates (with respect to the chosen basis). The following proposition will make it more clear.

Proposition 2.3. *Let V be a vector space and $v_1, \cdots, v_n \in V$. The following conditions are equivalent:*

(i) *$\{v_1, \cdots, v_n\}$ is a basis of V;*
(ii) *every vector $v \in V$ can be written uniquely as a linear combination of $v_1, \cdots, v_n$.*

In other words, the above proposition shows that there exists a bijection

$$V \to \mathbb{R}^n,$$
$$v \mapsto (c_1, ..., c_n)$$

such that

$$v = \sum_{i=1}^{n} c_i \cdot v_i.$$

This bijection is clearly dependent on the choice of the basis $\{v_1, \cdots, v_n\}$ of V and allows to associate to every vector an n-tuple of real numbers, that we will call *coordinates*.

Definition 2.5. *Let V a vector space and $\{v_1, \cdots, v_n\}$ a basis. For any $v = \sum_{i=1}^n c_i \cdot v_i$, the n-tuple $(c_1, \cdots, c_n) \in \mathbb{R}^n$ is called* **coordinates** *of v with respect to the basis $\{v_1, \cdots, v_n\}$.*

Remark 2.3. *In the case $V = \mathbb{R}^n$, one can check that the vectors:*

$$e_1 = (1, 0, 0, \cdots, 0),\ e_2 = (0, 1, 0, \cdots, 0),\ \cdots, e_n = (0, 0, 0, \cdots, 0, 1)$$

forms a basis. In this case, the coordinates of a vector $v = (v_1, \cdots, v_n)$ with respect to this basis are exactly its n components. This fact makes this basis the simplest possible, and in some sense the most natural among the other basis. For this reason this basis is called **canonical basis** *or* **standard basis** *of $\mathbb{R}^n$.*

3. Dimension of a Vector Space

In this section, we recall the following four important results about vector spaces:

- Every basis consists of the same number of vectors (theorem of dimension);
- it is possible to complete any set of linearly independent vectors, in order to obtain a basis for the space (theorem of completition);
- conversely, given any set of generators, it is possible to extract a basis for the space (theorem of reduction);
- the dimensions of two vector subspaces are related, via a precise formula, to the dimensions of the sum and the intersection spaces (Grassmann's dimension formula).

In other words, the number of vectors in a basis is a property of the space itself. Moreover, a basis is a maximal independent set (it cannot be made larger without losing the linear independence of the vectors) and a minimal spanning set (it cannot be made smaller and still span the space).

Theorem 3.1 (Theorem of dimension)**.** *Let V be a vector space and $\{v_1, \cdots, v_n\}$ a basis. Every other basis of V consists of n vectors.*

Definition 3.1. *Let V be a vector space. The number of vectors in a basis (and consequently in any other basis) is called the* **dimension** *of V and is denoted $\dim(V)$. If the number of vectors in a basis is finite, we say V is* **finite-dimensional***, otherwise, V is* **infinite dimensional**.

Example 3.1.

(1) $\dim(\mathbb{R}^n) = n$*; and* $\{e_1, \cdots, e_n\}$ *is a basis.*
(2) *the trivial subspace $Z = \{0\}$, has no basis. In fact, 0 is its only vector. We set* $\dim(Z) = 0$.
(3) *If $V = \langle v \rangle$, with $v \neq 0$, then* $\dim(V) = 1$ *(in fact $\{v\}$ is a basis).*

Theorem 3.2 (Theorem of completition)**.** *Let V be a vector space of dimension n and $w_1, \cdots, w_p \in V$ be p linearly independent vectors, with $p < n$. Then, there exist $n - p$ vectors $w_{p+1}, \cdots, w_n \in V$ such that $\{w_1, \cdots, w_n\}$ is a basis of V.*

Roughly speaking, any linearly independent set in V can be extended to a basis, by adding more vectors if necessary.

The next theorem says that any spanning set in V can be reduced to a basis, by discarding vectors if necessary.

Theorem 3.3 (Theorem of reduction)**.** *Let V be an n-dimensional vector space and $\{v_1, \cdots, v_m\}$ a spanning set of V. There exist n vectors*

$$v_{i_1}, \cdots, v_{i_n} \in \{v_1, \cdots, v_m\}$$

that form a basis of V.

Definition 3.2. *Let* $\dim(V) = n$ *and let $w_1, \cdots, w_p \in V$. We define the* **rank** *of $w_1, \cdots, w_p$ (denoted by* $\operatorname{rank}(w_1, \cdots, w_p)$*) as the dimension of the vector subspace spanned by these vectors; i.e.,*

$$\operatorname{rank}(w_1, \cdots, w_p) = \dim(\langle w_1, \cdots, w_p \rangle).$$

Remark 3.1. *Note that* $\operatorname{rank}(w_1, \cdots, w_p)$ *is the maximum number of linearly independent vectors among $w_1, \cdots, w_p$. Obviously,*

$$\operatorname{rank}(w_1, \cdots, w_p) \leq \min\{p, \dim(V)\}.$$

Finally, the following theorem shows how the dimensions of two vector subspaces are related to the dimensions of the sum and the intersection spaces; this is known as Grassmann's dimension formula.

Theorem 3.4 (Grassmann's dimension formula)**.** *Let V be a vector space and W_1, W_2 be two vector subspaces of V. We have:*

$$\dim(W_1) + \dim(W_2) = \dim(W_1 + W_2) + \dim(W_1 \cap W_2).$$

4. Useful Examples of Infinite-dimensional Vector Spaces

Example 4.1 (The sequence space l^∞)**.** *Let l^∞ be the set whose elements are of the form $\vec{x} = (x_1, x_2, \cdots, x_n, \cdots)$ where $x_i, i = 1, 2, \cdots$ are complex and $\{x_i\}$ are bounded; i.e.,*

$$\max_i |x_i| < \infty.$$

We define the two operations:

$$\begin{aligned} \vec{x} + \vec{y} &= (x_1 + y_1, x_2 + y_2, \cdots), \\ \alpha\vec{x} &= (\alpha x_1, \alpha x_2, \cdots), \end{aligned} \tag{4.1}$$

and check the properties of vector spaces. Note that $\vec{0} = (0, 0, \cdots)$ is the zero vector.

Example 4.2 (Sequence spaces l^p)**.** *Let l^p be the set consisting of elements $\vec{x} = (x_1, x_2, \cdots, x_n, \cdots)$, where $\sum_{i=1}^\infty |x_i|^p < \infty$. Note that if $p = 1$ then this is the set of all absolutely convergent sequences. The two operations are given in Eq. (4.1). To see that l^p is a vector space, we need to check that these operations produce vectors that are again in the space l^p. It is easy to check the eight properties for $p = 1$ to conclude that l^1 is a vector space. For example,*

$$\begin{aligned} &\sum_{i=1}^\infty |x_i + y_i| \le \sum_{i=1}^\infty (|x_i| + |y_i|) = \sum_{i=1}^\infty |x_i| + \sum_{i=1}^\infty |y_i| < \infty, \\ &\sum_{i=1}^\infty |\alpha x_i| < \infty, \quad \textit{for all} \quad \alpha \in \mathbb{R}. \end{aligned}$$

However, if $1 < p < \infty$, it is not obvious that $\vec{x} + \vec{y} \in l^p$ if $\vec{x}, \vec{y} \in l^p$; that is, whether $\sum_{i=1}^\infty |x_i + y_i|^p < \infty$, if $\sum_{i=1}^\infty |x_i|^p < \infty$ and $\sum_{i=1}^\infty |y_i|^p < \infty$. This works out because we have the following Minkowski inequality:

$$\left(\sum_{i=1}^\infty |x_i + y_i|^p\right)^{1/p} \le \left(\sum_{i=1}^\infty |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^\infty |y_i|^p\right)^{1/p}.$$

To summarize, l^p is a vector space for all $p \in [1, \infty]$.

Note that the dimension of l^p is infinite. To see this, let

$$\vec{e}_j = (\underbrace{0, \cdots, 0}_{j-1 \ \textit{entries}}, 1, 0, \cdots). \tag{4.2}$$

Obviously, $\vec{e}_j \in l^p$ and $\{\vec{e}_j\}_{j=1}^\infty$ is linearly independent. If l^p were finite dimensional, then there would be a finite basis $\{v_1, \cdots, v_n\}$ in l^p and any set of vectors with more than n vectors would be linearly dependent. This contradicts the linear independence of $\{\vec{e}_j\}_{j=1}^\infty$.

The following are some typical function spaces.

Example 4.3. *Let $\Omega \subset \mathbb{R}^n$ be a given subset of $\mathbb{R}^n$ for some integer n. Let $C(\Omega; \mathbb{R})$ be the set of all continuous functions from Ω to $\mathbb{R}$, together with the usual rules for addition and scalar multiplication of functions. It is easy to show that the sum of two continuous functions is continuous, and the product of a scalar and a continuous function is also continuous. (We say that addition and scalar multiplication are "well-defined" or "closed"). Note also that the zero function is continuous; $f(x) \equiv 0 \in C(\Omega; \mathbb{R})$. The eight properties of a vector space are satisfied. Usually, we denote $C(\Omega) = C(\Omega; \mathbb{R})$.*

Example 4.4. *$C(\Omega;\mathbb{C})$ is the set of all continuous functions from Ω to $\mathbb{C}$, which is a vector space.*

Example 4.5. *$C^k(\Omega;\mathbb{R}) = C^k(\Omega)$ and $C^k(\Omega;\mathbb{C})$ are functions with continuous derivatives up to k^{th} order.*

Note that all the above function spaces are infinite dimensional. For instance, for the function space $C(\mathbb{R};\mathbb{R})$, the set $\{1, x, x^2, \cdots, x^n, \cdots\}$ is linearly independent, and is a subset of the set of continuous functions.

Examples of subspaces. Recall that a subspace of a vector space V is a subset of V that is closed under addition and scalar multiplication; cf. Definition 1.2.

Example 4.6. *Let $p \in [1, +\infty)$ and S be the set of sequences $(0, x_1, \cdots, x_n, \cdots)$ such that*

$$\sum_{i=1}^{\infty} |x_i|^p < \infty.$$

Then S is a subset of l^p. It is easy to check the addition and scalar multiplication are closed in S; therefore, S is a subspace of l^p. However, the set S_1 defined by

$$S_1 := \{(1, x_1, \cdots, x_n, \cdots) \mid \sum_{i=1}^{\infty} |x_i|^p < \infty\}$$

is not a subspace of l^p because the zero element $(0, \cdots, 0, \cdots)$ is not in S_1.

Example 4.7. *Let H be the set of solutions of the equation*

$$y'' + p(t)y' + q(t)y = 0, \tag{4.3}$$

where $p(t)$ and $q(t)$ are continuous in $\mathbb{R}$. Then H is a subspace of $C^2(\mathbb{R})$.

Check addition: suppose $y_1(t)$ and $y_2(t)$ are two solutions of Eq. (4.3). Then,

$$\begin{aligned}
&(y_1 + y_2)'' + p(t)(y_1 + y_2)' + q(t)(y_1 + y_2) \\
&= y_1'' + y_2'' + p(t)(y_1' + y_2') + q(t)y_1 + q(t)y_2 \\
&= y_1'' + p(t)y_1' + q(t)y_1 + y_2'' + p(t)y_2' + q(t)y_2 \\
&= 0,
\end{aligned}$$

which implies

$$y_1(t) + y_2(t) \in H.$$

Check scalar multiplication: $\alpha y_1 \in H$. Therefore, H is a subspace of $C^2(\mathbb{R})$.

Example 4.8. *Let Ω be an open set with a smooth boundary. Define*

$$\widetilde{L}^2(\Omega;\mathbb{R}) = \{u \in C(\Omega;\mathbb{R}) \,\Big|\, \int_\Omega |u(x)|^2 \,\mathrm{d}x < +\infty\}. \tag{4.4}$$

Then $\widetilde{L}^2(\Omega;\mathbb{R})$ is a subspace of $C(\Omega;\mathbb{R})$.

Remark 4.1. *In Example 4.8 above, functions in $\widetilde{L}^2(\Omega;\mathbb{R})$ are called square integrable functions. Note that the domain Ω is not required to be bounded. Actually, if Ω is bounded, then any continuous functions that can be continuously extended to the boundary $\partial\Omega$ are square integrable.*

5. Inner Product Spaces

Note that the definition of a vector space gives an algebraic structure for vectors (i.e., addition and scalar multiplication). In the following, we shall build furthermore a geometric structure for vectors, such as length, distance, angle, etc.

Let us first introduce the notion of *inner product*, which is a generalization of the dot product on the Euclidean space $\mathbb{R}^n$.

Definition 5.1. *Let V be a real vector space. Consider a function from $V \times V$ to $\mathbb{R}$ that maps any pair of $u, v \in V$ to a real number. The function is called an* **inner product** *denoted by $\langle \cdot, \cdot \rangle$, (i.e., $\langle u, v \rangle$ is a real number) if for any $u, v, w \in V$ and $\alpha \in \mathbb{R}$ it satisfies*

(i) $\langle u, v \rangle = \langle v, u \rangle$,
(ii) $\langle u + w, v \rangle = \langle u, v \rangle + \langle w, v \rangle$, *(additivity)*
(iii) $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$,
(iv) $\langle u, u \rangle \geq 0$ *(non-negativity), and* $\langle u, u \rangle = 0$ *iff* $u = 0$ *(non-degeneracy).*

If V is a complex vector space, then an inner product on V is a function from $V \times V$ to $\mathbb{C}$ that satisfies (ii) *and* (iv) *above as well as the following complex conjugate properties* (i′) *and* (iii′) *in place of* (i) *and* (iii) *above:*

(i′) $\langle u, v \rangle = \overline{\langle v, u \rangle}$,
(iii′) $\langle u, \alpha v \rangle = \overline{\alpha} \langle u, v \rangle$.

Let us define next the notion of *lengths*, which is called a **norm** on the given vector space.

Definition 5.2. *Consider a function from a vector space V to $\mathbb{R}$. The function is called a* **norm***, denoted by $\| \cdot \|$ if it satisfies*

(i) $\|u\| \geq 0$ *and* $\|u\| = 0$ *iff* $u = 0$,
(ii) $\|\alpha u\| = |\alpha| \|u\|$, *for all* $\alpha \in \mathbb{R}$,
(iii) $\|u + v\| \leq \|u\| + \|v\|$.

A vector space on which a norm is defined is then called a **normed space** or **normed vector space**.

Theorem 5.1. *Given an inner product $\langle \cdot, \cdot \rangle$ on a vector space V, the function defined by $\|u\| = \sqrt{\langle u, u \rangle}$ for all $u \in V$ is a norm.*

Theorem 5.2 (Cauchy-Schwarz inequality)**.** *Given an inner product $\langle \cdot, \cdot \rangle$ on a vector space V, for any $u, v \in V$, it holds that*

$$|\langle u, v \rangle| \leq \|u\| \|v\|. \tag{5.1}$$

Example 5.1. *In $\mathbb{R}^n$, the dot product is an inner product:*

$$\langle u, v \rangle = u \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

The corresponding Euclidean norm is given by

$$\|u\| = \sqrt{\langle u, u \rangle} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}.$$

Example 5.2. *Let $\widetilde{L}^2(\Omega; \mathbb{R})$ be the subspace of $C(\Omega; \mathbb{R})$ defined in Example 4.8. Define*

$$\langle f, g \rangle = \int_\Omega f(x) g(x) \, \mathrm{d}x, \qquad \forall\, f, g \in \widetilde{L}^2(\Omega; \mathbb{R}). \tag{5.2}$$

Then $\langle\cdot,\cdot\rangle$ *is an inner product on* $\widetilde{L}^2(\Omega;\mathbb{R})$. *The corresponding norm is given by*

$$\|f\|_2 = \Big(\int_\Omega (f(x))^2\,\mathrm{d}x\Big)^{1/2}, \qquad \forall\, f\in \widetilde{L}^2(\Omega;\mathbb{R}) \tag{5.3}$$

Remark 5.1. *If we define* $L^2(\Omega;\mathbb{R})$ *to be the set of measurable functions from* Ω *to* $\mathbb{R}$ *that is square integrable (but not necessarily continuous). Then one can show that* $L^2(\Omega;\mathbb{R})$ *is a vector space and* $\langle f,g\rangle = \int_\Omega f(x)g(x)\,\mathrm{d}x$ *defines an inner product on* $L^2(\Omega;\mathbb{R})$. *Apparently,* $\widetilde{L}^2(\Omega;\mathbb{R})$ *is a subset of* $L^2(\Omega;\mathbb{R})$. *If you think about a Fourier series for a periodic square wave, you know that a sequence of continuous bounded functions (partial sums of the Fourier series) can converge to a discontinuous function (not a member of* $\widetilde{L}^2(\Omega;\mathbb{R})$*). If you allow such limits into the space, then we get* $L^2(\Omega;\mathbb{R})$. $L^2(\Omega;\mathbb{R})$ *is the completion of* $\widetilde{L}^2(\Omega;\mathbb{R})$.

5.1. **Orthogonal bases.** Let us first introduce the notion of *angle*. Recall the following property of the dot product on $\mathbb{R}^n$:

$$u\cdot v = \|u\|\|v\|\cos(\theta), \tag{5.4}$$

where $\theta\in[0,\pi]$ is the angle formed by the two vectors u and v. This can be generalized to any inner product space in a natural way by defining

$$\cos\theta = \frac{\langle u,v\rangle}{\|u\|\|v\|}, \tag{5.5}$$

where u and v are two vectors in a given inner product space V. Note that it follows from the Cauchy-Schwarz inequality (5.1) that the ratio $\frac{\langle u,v\rangle}{\|u\|\|v\|}$ is indeed between -1 and 1. The $\theta\in[0,\pi]$ that satisfies Eq (5.5) is called the **angle** between u and v.

Definition 5.3. *Let* V *be a vector space with an inner product* $\langle\cdot,\cdot\rangle$. *Two vector* u *and* v *in* V *are called* **orthogonal** *(or* **perpendicular***) if*

$$\langle u,v\rangle = 0.$$

A set of vectors $\{v_1,\cdots,v_n\}$ *in* V *are* orthogonal *if*

$$\langle v_i,v_j\rangle = 0 \quad \text{for any} \quad i\neq j.$$

An orthogonal set $\{v_1,\cdots,v_n\}$ *is called* orthonormal *if the set is orthogonal and* $\langle v_i,v_i\rangle = 1$ *for all* $i=1,\cdots,n$*; namely,* $\|v_i\|=1$.

If a set $\mathfrak{B}$ *of vectors in* V *is both an orthogonal (resp. orthonormal) set and a basis of* V, *then* $\mathfrak{B}$ *is called an orthogonal (resp. orthonormal) basis.*

Example 5.3. *The set of canonical basis* $\{e_i \mid i=1,\cdots,n\}$ *forms an orthonormal basis of* $\mathbb{R}^n$.

We have the following generalization of the Pythagorean theorem to high-dimensional case.

Theorem 5.3 (Pythagorean Theorem)**.** *Let* V *be a finite-dimensional inner product space with dimension* n*; and* $\mathfrak{B} := \{e_1,e_2,...,e_n\}$ *be an orthonormal basis of* V. *Then, for any* $u\in V$, *it holds that*

$$u = \langle u,e_1\rangle e_1 + \langle u,e_2\rangle e_2 + \cdots + \langle u,e_n\rangle e_n = \sum_{i=1}^n \langle u,e_i\rangle e_i, \tag{5.6}$$

and

$$\|u\|^2 = |\langle u,e_1\rangle|^2 + |\langle u,e_2\rangle|^2 + \cdots + |\langle u,e_n\rangle|^2 = \sum_{i=1}^n |\langle u,e_i\rangle|^2. \tag{5.7}$$

Remark 5.2. *The above results can be directly generalized to the case of infinite-dimensional inner product space. Of course, care needs to be taken when dealing with infinite summation; and it turns out that there are infinite-dimensional inner product spaces that do not have an orthonormal basis. However, every complete inner product space (i.e., a Hilbert space) has an orthonormal basis.*

Before we state the infinite-dimensional version of the above Pythagorean theorem, we need to introduce the following concepts:

Definition 5.4. *Given a norm* $\|\cdot\|$ *on a vector space* V*, a sequence of vectors* $\{v_1, v_2, \cdots\}$ *converges to a vector* $v \in V$ *if*

$$\lim_{n\to\infty} \|v_n - v\| = 0. \tag{5.8}$$

Definition 5.5 (Complete orthonormal set)**.** *Let* V *be an infinite-dimensional inner product space and* $\mathfrak{B} := \{e_i \mid i \in \mathbb{N}\}$ *be an orthonormal set in* V*.* $\mathfrak{B}$ *is complete if for every* $v \in V$*, it holds that*

$$\lim_{n\to\infty} \|v - \sum_{i=1}^{n} \langle v, e_i\rangle e_i\| = 0. \tag{5.9}$$

If Eq (5.9) *holds, we write*

$$v = \sum_{i=1}^{\infty} \langle v, e_i\rangle e_i. \tag{5.10}$$

Of course, a complete orthonormal set forms an orthonormal basis.

Theorem 5.4 (Pythagorean Theorem, infinite-dimensional version)**.** *Let* V *be an infinite-dimensional inner product space, and* $\mathfrak{B} := \{e_1, e_2, ...\}$ *be an orthonormal basis of* V*. Then, for any* $u \in V$*, it holds that*

$$u = \sum_{i=1}^{\infty} \langle u, e_i\rangle e_i, \tag{5.11}$$

and

$$\|u\|^2 = \sum_{i=1}^{\infty} |\langle u, e_i\rangle|^2. \tag{5.12}$$

Example 5.4. *Consider a function* f *on the interval* $[0, L]$ *and the corresponding Fourier sine series:*

$$f(x) \sim \sum_{n=1}^{\infty} a_n \sin(\frac{n\pi x}{L}). \tag{5.13}$$

Let us denote

$$\phi_n(x) = \sin(\frac{n\pi x}{L}). \tag{5.14}$$

We can check by using $\sin(x)\sin(y) = \frac{1}{2}(\cos(x-y) - \cos(x+y))$ *that*

$$\langle \phi_n, \phi_m\rangle := \int_0^L \phi_n(x)\phi_m(x)\,\mathrm{d}x = \begin{cases} 0, & \text{if } m \neq n, \\ \frac{L}{2}, & \text{if } m = n. \end{cases} \tag{5.15}$$

Then by defining

$$e_n = \sqrt{\frac{2}{L}} \sin(\frac{n\pi x}{L}), \tag{5.16}$$

we obtain an orthonormal set $\{e_n \mid n \in \mathbb{N}\}$. In fact $\{e_n \mid n \in \mathbb{N}\}$ is also complete in $L^2([0,L])$. Using then the fact that $L^2([0,L])$ is a complete normed space under the norm induced by the L^2-inner product, and assuming that $f \in L^2([0,L])$, one obtains that

$$f(x) = \sum_{n=1}^{\infty} \langle f, e_n \rangle e_n, \tag{5.17}$$

where

$$\langle f, e_n \rangle = \int_0^L f(x) e_n(x) \, \mathrm{d}x. \tag{5.18}$$

Now, using Eq (5.17) and the definition of e_n given in Eq (5.16), we obtain the following expression for the coefficients a_n appeared in the Fourier series (5.13):

$$a_n = \frac{2}{L} \int_0^L f(x) \sin(\frac{n\pi x}{L}) \, \mathrm{d}x. \tag{5.19}$$

6. Sturm-Liouville Boundary Value Problem

In this section, we record some useful results related to the theory of Sturm-Liouville boundary value problems.

Recall that a **regular** Sturm-Liouville boundary value problem on a bounded interval $[a,b] \subset \mathbb{R}$ is of the following form:

$$\begin{aligned} -(p(x)u'(x))' + q(x)u(x) &= \lambda w(x) u(x), \quad x \in (a,b), \\ \alpha_1 u(a) + \alpha_2 u'(a) &= 0, \\ \beta_1 u(b) + \beta_2 u'(b) &= 0, \end{aligned} \tag{6.1}$$

where p, q, w are given continuous functions on $[a,b]$ with $p > 0$ and $w > 0$, p' is also assumed to be continuous, $\alpha_1^2 + \alpha_2^2 \neq 0$, and $\beta_1^2 + \beta_2^2 \neq 0$.

Example 6.1. *Note that the eigenvalue problem*

$$\begin{aligned} u''(x) &= \lambda u, \qquad x \in (0,L), \\ u(0) &= 0, \\ u(L) &= 0, \end{aligned} \tag{6.2}$$

fits the above general form (6.1).

Definition 6.1. *A series of functions $\sum_{n=1}^{\infty} u_n(x)$ defined on $\mathbb{R}$* **converges** *to (or has a sum) $S(x)$ if its sequence of partial sums $S_n(x) = \sum_{n=1}^{\infty} u_n(x)$ converges to $S(x)$. That is, for any given $\epsilon > 0$, there is an integer $N(\epsilon, x)$ depending on ϵ and x such that:*

$$|S_n(x) - S(x)| < \epsilon, \qquad \text{for all } n > N(\epsilon, x). \tag{6.3}$$

If for a given interval I, $N(\epsilon, x)$ can be chosen to be independent of $x \in I$, then $\sum_{n=1}^{\infty} u_n(x)$ is said to be **uniformly convergent** *to $S(x)$ in the interval I.*

Theorem 6.1 (Weierstrass M-test)**.** *Consider a series of functions $\sum_{n=1}^{\infty} u_n(x)$ defined on $\mathbb{R}$. For a given interval $I \subset \mathbb{R}$, if there is a convergent series of positive constants $\sum_{n=1}^{\infty} M_n$ such that $|u_n(x)| \leq M_n$ for each n and all $x \in I$, then $\sum_{n=1}^{\infty} u_n(x)$ is uniformly convergent in I.*

The following are some nice properties of uniformly convergent series of continuous functions:

(i) A uniformly convergent series of continuous functions must converge to a continuous function.

(ii) If a series $\sum_{n=1}^{\infty} u_n(x)$ of continuous functions converges uniformly to $S(x)$ for $x \in [a,b]$, then

$$\int_b^a S(x)\,\mathrm{d}x = \sum_{n=1}^{\infty} \int_a^b u_n(x)\,\mathrm{d}x. \tag{6.4}$$

(iii) Assume $\sum_{n=1}^{\infty} u_n(x) = S(x)$ for $x \in [a,b]$. If $u_n'(x)$ is continuous and $\sum_{n=1}^{\infty} u_n'(x)$ is uniformly convergent, then

$$S'(x) = \sum_{n=1}^{\infty} u_n'(x), \qquad \text{for all } x \in [a,b]. \tag{6.5}$$

Theorem 6.2 (Properties of eigenfunction expansions)**.** *Let $\phi_n(x)$, $n = 1, 2, \cdots$, be the eigenfunctions of a regular Sturm-Liouville boundary value problem of the form (6.1). For a given function $f \in L_w^2([a,b])$, if both $f(x)$ and $f'(x)$ are piecewise continuous, then*

$$\sum_{n=1}^{\infty} c_n \phi_n(x) = \frac{f(x+) + f(x-)}{2}, \quad \textit{for all} \quad x \in (a,b), \tag{6.6}$$

where

$$\begin{aligned} f(x+) &:= \lim_{y \to x^+} f(y), \\ f(x-) &:= \lim_{y \to x^-} f(y), \end{aligned} \tag{6.7}$$

and the coefficient c_n is given by

$$c_n = \int_a^b f(x)\phi_n(x)w(x)\,\mathrm{d}x, \qquad \textit{for all} \quad n \in \mathbb{N}. \tag{6.8}$$

Moreover, if f also satisfies the given boundary conditions in (6.1), then it holds that

$$\sum_{n=1}^{\infty} c_n \phi_n(a) = f(a) \quad \textit{and} \quad \sum_{n=1}^{\infty} c_n \phi_n(b) = f(b). \tag{6.9}$$

The case of periodic boundary conditions. In the remaining part of this section, we consider another useful setting concerning periodic boundary conditions that cannot be put into the general form given by (6.1).

Let us consider

$$\begin{aligned} -(p(x)u'(x))' + q(x)u(x) &= \lambda w(x)u(x), \quad x \in (a,b), \\ u(a) &= u(b), \\ u'(a) &= u'(b), \end{aligned} \tag{6.10}$$

where we assume $p(a) = p(b)$ besides the other conditions on p, q, w given before.

Note that (6.10) is **not** a Sturm-Liouville boundary value problem. Unlike the boundary conditions in (6.1) where the condition at the endpoint $x = a$ is separated from the condition at the endpoint $x = b$. Here, the boundary conditions mix the information at the two endpoints. The boundary conditions in (6.10) are called periodic boundary conditions.

All the theorems above regarding regular Sturm-Liouville boundary value problems still hold for the periodic case (6.10) except that the geometric multiplicity of the eigenvalues in the periodic case is no longer 1 except for the smallest eigenvalue.

Appendix A. Complexification of a real inner product space

The complexification of a real inner product space V is defined to be

$$\widetilde{V} := \{u + iv : u, v \in V\} \tag{A.1}$$

endowed with the inner product

$$\langle \widetilde{u}, \widetilde{v} \rangle_{\widetilde{V}} := (\langle u_R, v_R \rangle_V + \langle u_I, v_I \rangle_V) + i(\langle u_I, v_R \rangle_V - \langle u_R, v_I \rangle_V), \tag{A.2}$$

where $\widetilde{u} = u_R + iu_I$ and $\widetilde{v} = v_R + iv_I$ with $u_R, u_I, v_R, v_I \in V$, and $\langle \cdot, \cdot \rangle_V$ denotes the inner product on V.

The complexification of an operator L on V is defined to be

$$\begin{aligned} \widetilde{L} \colon \widetilde{V} &\to \widetilde{V}, \\ u + iv &\mapsto L(u) + iL(v). \end{aligned} \tag{A.3}$$