Applied Linear Algebra

Contents

List of Symbols

$\forall$	For all/any
$\exists$	For some / There exists
$a \in S$	a belongs to / is an element of set S
$\cup$, $\cap$	Union and intersection (respectively) of sets
$\varnothing$	Empty set $\{\}$
$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$	Sets of natural numbers, integers, rational numbers, real numbers, complex numbers (respectively)
$\mathbb{N}_0$	Set of non-negative integers $\{0, 1, 2, \ldots\}$
$A \subseteq B$	A is a subset of set B
$\lvert S \rvert$	Cardinality (number of elements) of set S
F^n	The vector space of n-tuples of F-elements, where F is a field and $n \in \mathbb{N}$
$F^{m \times n}$	The vector space of $m \times n$ matrices over F, where F is a field and $m, n \in \mathbb{N}$
$\mathcal{C}(A)$, $\mathcal{R}(A)$	The column space and row space of a matrix A

1 Introduction

Linear algebra is built around the notion of linearity, and the operation of forming linear combinations. To formally state what these terms mean, we need to define vector spaces. To do this in the most general manner possible, we define a vector space axiomatically – i.e., by specifying properties that any vector space should satisfy, rather than by specifying what structure a vector should have. So we define vectors neither as quantities with direction and magnitude (which, as it turns out, are defined not in vector spaces but in inner product spaces and normed linear spaces), nor as n-tuples of real numbers, but instead as elements of a vector space as defined by the axioms.

This reversal in the order of definitions by means of abstraction is a staple of higher mathematics, as it results in greater generality. All the familiar examples of vectors become special cases of our more general definition. But we also obtain new examples of vector spaces (e.g., spaces of functions) that would not have fit into the earlier definitions.

2 Vector Spaces

Definition 2.1. An Abelian group (or commutative group) is an ordered pair $(A, *)$, where A is a set and $*\colon A \times A \longrightarrow A$ is a binary operation on A satisfying the following conditions:

1. Associativity: $\forall a, b, c \in A$, $a * (b * c) = (a * b) * c$.
2. Existence of identity: $\exists e \in A$, $\forall a \in A$, $a * e = e * a = a$. The element e is called the identity element of the group.
3. Existence of inverses: $\forall a \in A$, $\exists b \in A$, $a * b = b * a = e$ (where e is the identity element given in 2. The element b is then called an inverse of a.
4. Commutativity: $\forall a, b \in A$, $a * b = b * a$.

Conditions 1–4 are called the axioms of Abelian groups. The first three axioms alone define a group and it is the fourth axiom of commutativity that makes the group Abelian.

Axiom 2 states that every element in the group has *an* inverse. But we can prove that such an inverse must be unique. Thus, every element $a \in A$ has a *unique* inverse, which we shall denote as a^{-1}.

Example 2.2. The sets of integers, rationals, reals, and complex numbers all form Abelian groups under their usual addition: $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$. In each case, the identity element is 0, and the inverse of x is its negative, $-x$.

Example 2.3. The sets of *non-zero* rationals, reals, and complex numbers form Abelian groups under multiplication: $(\mathbb{Q}-\{0\},\times)$, $(\mathbb{R}-\{0\},\times)$, $(\mathbb{C}-\{0\},\times)$. Note that the non-zero integers do *not* form a group under multiplication (Why?).

Example 2.4. Let $\mathbb{Z}_n = \{0,1,\ldots,n-1\}$, and let $+_n$. That is, $a +_n b$ is the remainder obtained when the integer sum $a+b$ is divided by n. Then $(\mathbb{Z}_n, +_n)$ is an Abelian group with exactly n elements.

Definition 2.5. A field is a triple $(F,+,\cdot)$, where $+\colon F\times F\to F$ and $\cdot\colon F\times F\to F$ are binary operations on F satisfying the following conditions:

1. $(F,+)$ is an Abelian group.
2. $(F-\{0\},\cdot)$ is an Abelian group, where 0 denotes the identity of the Abelian group $(F,+)$.
3. Distributivity of $\cdot$ over $+$: $\forall a,b,c\in F$, $a\cdot(b+c)=a\cdot b+a\cdot c$.

The operations $+$ and $\cdot$ are respectively called the addition and multiplication of the field F. Thus, the first axiom of fields given above states that the elements of the field form an Abelian group under addition. We denote its identity element (called the additive identity or zero of the field) by 0, and the inverse of any element a by $-a$ (called the additive inverse of a). The second axiom states that the non-zero field elements form an Abelian group under multiplication. We denote the multiplicative identity[1] by 1 and the multiplicative inverse of an element $a\neq 0$ by a^{-1} or $\frac{1}{a}$. If a and $b\neq 0$ are two elements of F, then we will write a/b or $\frac{a}{b}$ to mean $a\cdot b^{-1}=b^{-1}a$. We usually drop the operator $\cdot$ and simply write ab to mean $a\cdot b$.

The axioms of a field are defined in such a way that they give rise to most of the familiar laws of arithmetic, such as the cancellation laws: $\forall a,b,c\in F$, $a+b=a+c \implies b=c$ and if $a\neq 0$, then $ab=ac \implies b=c$.

Example 2.6. $(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$, and $(\mathbb{C},+,\cdot)$, are fields. Examples 2.2 and 2.3 state that in each of these cases, the set forms an Abelian group under addition and its non-zero elements form an Abelian group under multiplication. Finally, we know that in each case, multiplication distributes over addition. Thus, all the field axioms are satisfied in each case. All three are examples of fields with infinitely many elements.

Example 2.7. Let $\mathbb{Z}_n$ and $+_n$ be as defined in Example 2.4, and let $\times_n$ be multiplication modulo n, defined similarly to $+_n$. We can show that $(\mathbb{Z}_n,+_n,\times_n)$ is a field if and only if n is a prime number (Exercise). Thus, there exists a finite field with n element for every prime number n.

[1] Exercise: According to the second axiom, $1\cdot a=a$ for all $a\neq 0$. Show that $1\cdot 0=0$ as well, and that $0\cdot a=0$ for all $a\in F$.

Definition 2.8. A vector space over a field F is a a triple $(V, +, \cdot)$, where $+ \colon V \times V \longrightarrow V$ is a binary operation on V, called vector addition, $\cdot \colon F \times V \longrightarrow V$ is a binary operation called scalar multiplication, satisfying the following conditions:

1. $(V, +)$ is an Abelian group.
2. $\forall v \in V$, $1 \cdot v = v$, where 1 is the multiplicative identity of F.
3. Scalar multiplication distributes over addition of field elements and vector addition: $\forall \alpha, \beta \in F$, $u, v \in V$,
$$\begin{aligned}(\alpha + \beta) \cdot v &= \alpha \cdot v + \beta \cdot v \\ \alpha \cdot (u + v) &= \alpha \cdot u + \alpha \cdot v\end{aligned}$$
4. Scalar multiplication associates with the field multiplication: $\forall \alpha, \beta \in F$, $v \in V$,
$$\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v.$$

The elements of V are called vectors, and those of F are called scalars. The identity element of vector addition is denoted by 0_V, or 0, and is called the zero vector – the context usually makes it clear whether by 0 we mean the zero vector or the scalar zero.

Theorem 2.9. *Let V be a vector space over a field F. Then*

1. $\forall v \in V$, $0 \cdot v = 0_V$ *and* $(-1) \cdot v = -v$.
2. $\forall \alpha \in F$, $\alpha \cdot 0_V = 0_V$
3. $\forall \alpha \in F$, $v \in V$, *if* $\alpha \cdot v = 0_V$ *and* $\alpha \neq 0$, *then* $v = 0_V$.

Proof. Let $v \in V$, $\alpha \in F$.

1. We have
$$\begin{aligned}0 \cdot v + 0 \cdot v &= (0 + 0) \cdot v && \text{by distributivity} \\ &= 0 \cdot v \\ &= 0 \cdot v + 0_V && \text{by identity law in } (V, +).\end{aligned}$$
Then by cancellation in the Abelian group $(V, +)$, $0 \cdot v = 0_V$.
Now,
$$\begin{aligned}v + (-1) \cdot v &= (1 + -1) \cdot v && \text{by distributivity, since } 1 \cdot v = v \\ &= 0 \cdot v \\ &= 0_V.\end{aligned}$$
Thus, $(-1) \cdot v$ is the additive inverse of v, which is $-v$.

2. Similarly, $\alpha \cdot 0_V = \alpha \cdot 0_V + \alpha \cdot 0_V \implies \alpha \cdot 0_V = 0_V$.
3. If $\alpha \neq 0$, then $\alpha \cdot v = 0_V \implies$ $\alpha^{-1} \cdot 0_V = \alpha^{-1} \cdot (\alpha \cdot v) = (\alpha^{-1}\alpha) \cdot v = 1 \cdot v = v \implies$ $0_V = v$.

□

Example 2.10 (n-Tuple Space)**.** For any field F and any fixed positive integer n, define

$$F^n = \{ (x_1, x_2, \dots, x_n) \mid x_1, x_2, \dots, x_n \in F \},$$

the set of all n-tuples of F-elements. Then $(F^n, +, \cdot)$ is a vector space over F itself, where $+$ denotes componentwise addition of n-tuples, and $\cdot$ denotes componentwise multiplication of an n-tuple by a field element (verify this). Explicitly, if $x = (x_1, \dots, x_n)$, and $y = (y_1, \dots, y_n)$, then we define $x + y$ to be the n-tuple whose i^{th} component is $x_i + y_i$, and for any $\alpha \in F$, we define $\alpha \cdot x$ to be the n-tuple whose i^{th} component is αx_i, for each $i = 1, \dots, n$. As a convention, if we say "the vector space F^n", we will mean the vector space of n-tuples of F-elements, over the field F, as defined here. Particular cases of interest are $\mathbb{R}^n$ and $\mathbb{C}^n$.

Example 2.11. For any field F and positive integers m and n, define $F^{m \times n}$ to be the set of all $m \times n$ matrices with entries from F, and let $+$ denote matrix addition and $\cdot$ denote multiplication of a matrix by an element of F. Then $(F^{m \times n}, +, \cdot)$ is a vector space over F. In particular, $F^{m \times 1}$ and $F^{1 \times n}$ are the vector spaces consisting of all column vectors of length m, and all row vectors of length n, respectively.

Example 2.12. Let V be the set of all functions from $\mathbb{R}$ to $\mathbb{R}$ (so any *element* of V is a function $f : \mathbb{R} \to \mathbb{R}$). Define a binary operator $+$ on V as follows: for any two functions $f, g \in V$, let $h = f + g$ be the function such that $h(x) = f(x) + g(x)$ for all $x \in \mathbb{R}$. Define $\cdot : \mathbb{R} \times V \to V$ as: for any $\alpha \in \mathbb{R}$ and $f \in V$, let $h = \alpha f$ be the function such that $h(x) = \alpha f(x)$. Then V is a vector space over $\mathbb{R}$.

Exercise 2.1. Generalise Example 2.12 by taking $V = F^S$ to be the set of all functions from S to F, where S is any fixed, non-empty set, and F is any field, and showing that this forms a vector space. In the particular case where $S = \{1, \dots, n\}$, justify the idea that F^S is essentially the same as the n-tuple space F^n.

Example 2.13 (Column Space)**.** Let F be a field, and let A be a given $m \times n$ matrix with entries from F. Let $x_1, \dots, x_n$ be the n different columns of A. Define

$$C(A) = \{ \alpha_1 x_1 + \cdots + \alpha_n x_n \mid \alpha_i \in F, i = 1, \dots, n \}.$$

$C(A)$ forms a vector space over F under the usual addition and scalar multiplication of column vectors (verify). This vector space is called the column space of the matrix A.

Exercise 2.2. Define the row space $\mathcal{R}(A)$ of the matrix A in a similar manner and verify that it is a vector space.

Exercise 2.3. Consider a system of linear equations

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + & a_{mn}x_n & = & b_m \end{array} \tag{1}$$

consisting of m equations in n unknowns $x_1, \ldots, x_n$. This can be written equivalently as a matrix equation $Ax = b$, where A, the coefficient matrix, is the matrix whose (i, j)-entry is the coefficient a_{ij} in the system (1); x, the vector of unknowns, is the column vector whose j^{th} entry is the unknown x_j; and b, the right hand side vector is the column vector whose i^{th} entry is the constant b_i in the i^{th} equation; for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

1. Describe the matrix product Ax in terms of the column vectors of A and components of x.
2. In terms of the column space $C(A)$ and vector b, when does the system (1) have a solution?

2.1 Linear Combinations

As stated in the introduction, linear algebra revolves around the operation of taking linear combinations. Having defined vector spaces formally, we are now in a position to formally define linear combinations as well.

Definition 2.14. Let V be a vector space over a field F, and let $v_1, \ldots, v_n \in V$. Then a linear combination of the vectors $v_1, \ldots, v_n$ is any vector of the form $\alpha_1 v_1 + \cdots + \alpha_n v_n$, where $\alpha_1, \ldots, \alpha_n \in F$.

If w is a given vector of V, and if there exist scalars $\alpha_1, \ldots, \alpha_n \in F$ such that $w = \alpha_1 v_1 + \cdots + \alpha_n v_n$, then we say that w is a linear combination of $v_1, \ldots, v_n$.

Remark. Our definition of $+$ is as a binary operation on V, and therefore we can only add *two* given vectors. However, an expression of the form $v_1 + \cdots + v_n$ is well defined, since it merely denotes repeated binary addition of vectors, and thanks to associativity, the order in which the additions are evaluated (i.e., parenthesization) does not affect the value of the sum. Note, however, than an *infinite* sum $v_1 + v_2 + \cdots$ is not defined, unless almost all[2] summands are zero, in which case we define the value of this infinite sum to be equal to the sum of the non-zero summands. Therefore, even if we consider an infinite set of vectors, we can only form linear combinations of finitely many of them, and in such cases we should be careful to mention that we are taking a *finite* linear combination.

3 Subspaces

A subspace of a vector space V is a subset $W \subseteq V$ such that W itself is a vector space under the vector addition and scalar multiplication operations that it inherits[3] from V.

Example 3.1. Take $V = \mathbb{R}^3$, the real 3-tuple space, and let

$$W = \{ (x, y, 0) \mid x, y \in \mathbb{R} \}$$

which is obviously a subset of V. Since $(x_1, y_1, 0) + (x_2, y_2, 0) = (x_1 + x_2, y_1 + y_2, 0) \in W$, the addition in V is a binary operator on W as well. Similarly, the scalar multiplication of V, when restricted to W, also works as a scalar multiplication on W, since $\alpha(x, y, 0) = (\alpha x, \alpha y, 0) \in W$. It is then clear that the vector space axioms hold for W with these operations, and hence W is also a vector space over $\mathbb{R}$. Therefore, W is a subspace of V.

When checking whether the vector space axioms hold for a subset W of a vector space V, we do not need to verify commutativity of $+$, distributivity of $\cdot$ over $+$, associativity of scalar multiplication with the field multiplication, or the invariance of vectors under scalar multiplication by 1. Since all elements of W are elements of V, these conditions must automatically by satisfied by them. Instead, we need to check whether

1. W is closed under $+$ and $\cdot$ (for this is necessary for the restrictions of these operations to W to work as vector addition and scalar multiplication of W itself), and
2. $(W, +)$ is a group – which is equivalent to checking if $0_V \in W$, and $\forall w \in W$, $-w \in W$.

The implication of these conditions is that *a non-empty subset of a vector space is a subspace precisely when it is closed under linear combinations*. In fact, we can simplify that statement even further, as the following theorem shows.

[2] The term almost all is used formally here, to mean "all but finitely many".

[3] If $(V, +, \cdot)$ is a vector space over a field F, and $W \subseteq V$, then W is a subspace of V if the restrictions of $+$ and $\cdot$ to W are binary operations with images contained in W, and if W forms a vector space over F under these operations.

Theorem 3.2. *A subset W of a vector space V over a field F is a subspace of V if and only if $W \neq \varnothing$ and $\forall \alpha \in F$, $\forall u, v \in W$, $\alpha u + v \in W$.*

Proof. If W is a subspace, then $0_V \in W$, and hence $W \neq \varnothing$, and for any $\alpha \in F$, $u, v \in W$, we have $\alpha u \in W$ (by closure of W under $\cdot$) and then $\alpha u + v \in W$ (by closure of W under vector addition).

Now to prove the converse, suppose that W is a non-empty subset of V such that $\forall \alpha \in F$, $u, v \in W$, $\alpha u + v \in W$. We must prove that W is closed under $+$ and $\cdot$, that $0_V \in W$, and that $\forall v \in W$, $-v \in W$.

Since $W \neq \varnothing$, $\exists v \in W$. Now, $-1 \in F$ and $v, v \in W$, and hence $(-1)v + v = -v + v = 0_V \in W$. For closure under $+$, observe that if $u, v \in W$, then $1 \in F \implies 1u + v = u + v \in W$. For closure under $\cdot$, observe that if $\alpha \in F$ and $v \in W$, then $v, 0_V \in W \implies \alpha v + 0_v = \alpha v \in W$. Finally, if $v \in V$, then by closure under $\cdot$ we have $(-1)v = -v \in W$. □

This theorem simplifies the task of proving that certain subsets of a vector space are subspaces. For instance, we can use it to prove that the intersection of two subspaces is again a subspace.

Exercise 3.1. Let W_1 and W_2 be subspaces of a vector space V. Prove that their intersection $W_1 \cap W_2$ is a subspace of V.

This immediately implies (by mathematical induction) that the intersection of a finite number of subspaces is also a subspace. But even more generally, the intersection of *any* collection of subspaces is a subspace.

Theorem 3.3. *Let V be a vector space, and let $\{ W_i \mid i \in I \}$ be a collection of subspaces of V, indexed[4] by a set I, and let*

$$W = \bigcap_{i \in I} W_i.$$

Then W is a subspace of V.

Proof. Since $0 \in W_i$ for all $i \in I$, we have $0 \in W$, so that $W \neq \varnothing$. Let $\alpha \in F$, and $u, v \in W$. Then for each $i \in I$, $u, v \in W_i$, which is a subspace, and hence by Theorem 3.2, $\alpha u + v \in W_i$. Therefore, $\alpha u + v \in W$, the intersection of all the W_is. Again by Theorem 3.2, these two observations imply that W is a subspace. □

[4] A finite collection of objects (e.g., an infinite collection of some subsets) can be easily represented as a list, say $A_1, A_2, \ldots, A_n$. But to denote a possibly infinite collection of objects, we need to specify an indexing set I, and then denote each element in the collection by A_i for some $i \in I$. The whole collection is then $\{ A_i \mid i \in I \}$. Now I could be finite, countably infinite, or uncountable infinite. For example, if $I = \{1, \ldots, n\}$, then we get the finite collection $\{A_1, \ldots, A_n\}$. On other hand, it's possible that $I = \mathbb{N}$. Then the collection is countably infinite. Or, say, $I = \mathbb{R}$, so that the collection is uncountably infinite.

But in general, the *union* of two subspaces is **not** a subspace, except in a trivial way.

Theorem 3.4. *If W_1 and W_2 are subspaces of a vector space V, then $W_1 \cup W_2$ is not a subspace of V unless $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.*

Proof. If $W_1 \subseteq W_2$, then $W_1 \cup W_2 = W_2$, which is given to be a subspace. The other case is similar.

Therefore, suppose that neither of the subspaces W_1 and W_2 is contained in the other. We will show that $W_1 \cup W_2$ is not a subspace, by showing that it is not closed under vector addition.

Since $W_1 \nsubseteq W_2$, $\exists w_1 \in W_1$, such that $w_1 \notin W_2$. Similarly, $\exists w_2 \in W_2$, $w_2 \notin W_1$.

Let $w = w_1 + w_2$. We claim that $w \notin W_1$ and $w \notin W_2$. For if $w \in W_1$, then $w_2 = w - w_1 \in W_1$ as well (since $w_1, w \in W_1$, a subspace), which contradicts the definition of w_2. Similarly, if $w \in W_2$, then we get the contradiction that $w_1 \in W_2$.

Thus, $w_1, w_2 \in W_1 \cup W_2$, but $w = w_1 + w_2 \notin W_1 \cup W_2$, implying that $W_1 \cup W_2$ is not a subspace. □

In some sense, then, the set-theoretic operation of union is not a suitable operation for use in vector space (except, of course, when dealing with subsets rather than subspaces). That is, we cannot use it for generating new, possibly larger subspaces from given ones[5]. We now define an operation that *does* accomplish this.

Definition 3.5. If $W_1, \ldots, W_k$ are subspaces of a vector space V, then the sum of $W_1, \ldots, W_k$ is the subspace $W = W_1 + \cdots + W_k$, defined by

$$W = \{\, w_1 + \cdots + w_k \mid w_i \in W_i, i = 1, \ldots, k \,\}.$$

The definition *states* that this sum is a subspace, but this is a claim that needs to be verified. We will see later, after defining the span of a subset, that the sum of subspaces is exactly the span of their union, and is therefore a subspace. Observe that $W_1 + \ldots + W_k$ contains each subspace W_i, $i = 1, \ldots, k$.

Example 3.6. Let $V = \mathbb{R}^3$, and define $X = \{\, (x, 0, 0) \mid x \in \mathbb{R} \,\}$, $Y = \left\{ (0, y, 0) \mid y \in \mathbb{R} \right\}$, $Z = \{\, (0, 0, z) \mid z \in \mathbb{R} \,\}$ and $W = \left\{ (x, y, 0) \mid x, y \in \mathbb{R} \right\}$. It is easily verified that X, Y, Z, and W are subspaces, and that $W = X + Y$, and $V = X + Y + Z = W + Z$. Also note that $V = X + W + Z = Y + W + Z$. But $X + W = Y + W = W$.

[5] We can similarly define the sum of subsets that are not necessary subspaces, but such a sum is not guaranteed to be a subspace.