Chapter 5

Linear Algebra

The guiding theme in linear algebra is the interplay between algebraic manipulations and geometric interpretations. This dual representation is what makes linear algebra a fruitful and fascinating field of mathematics.

By the scalars, ($\mathbb{F}$), we shall mean either the field of real numbers or the field of complex numbers. A **linear space** (or **vector space**) X over $\mathbb{F}$ is, to begin with, a set X together with a binary operation, vector addition ($+$), on $X \times X$. To every pair of elements (vectors) $\mathbf{x}_i, \mathbf{x}_j \in X$, there corresponds a vector $\mathbf{x}_i + \mathbf{x}_j \in X$ (the sum of $\mathbf{x}_i$ and $\mathbf{x}_j$), such that

$$\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}_2 + \mathbf{x}_1, \qquad \mathbf{x}_1, \mathbf{x}_2 \in X,$$

$$(\mathbf{x}_1 + \mathbf{x}_2) + \mathbf{x}_3 = \mathbf{x}_1 + (\mathbf{x}_2 + \mathbf{x}_3), \qquad \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \in X,$$

$$\exists !\, \mathbf{0} \in X \text{ such that } \mathbf{0} + \mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{x}, \qquad \mathbf{x} \in X,$$

$$\forall\, \mathbf{x} \in X, \exists !\, (-\mathbf{x}) \in X \text{ such that } \mathbf{x} + (-\mathbf{x}) = \mathbf{0}.$$

Furthermore a vector space is by definition equipped with another operation, scalar multiplication, $m : \mathbb{F} \times X \to X$, which we shall denote by $m(\lambda, \mathbf{x}) = \lambda\mathbf{x}$, $\lambda \in \mathbb{F}$, $\mathbf{x} \in X$, such that

$$\lambda(\mathbf{x}_1 + \mathbf{x}_2) = \lambda\mathbf{x}_1 + \lambda\mathbf{x}_2, \qquad \lambda \in \mathbb{F}, \quad \mathbf{x}_1, \mathbf{x}_2 \in X,$$

$$(\lambda_1 + \lambda_2)\mathbf{x} = \lambda_1\mathbf{x} + \lambda_2\mathbf{x}, \qquad \lambda_1, \lambda_2 \in \mathbb{F}, \quad \mathbf{x} \in X,$$

$$(\lambda_1\lambda_2)\mathbf{x} = \lambda_1(\lambda_2\mathbf{x}), \qquad \lambda_1, \lambda_2 \in \mathbb{F}, \quad \mathbf{x} \in X,$$

$$1\mathbf{x} = \mathbf{x}, \qquad \mathbf{x} \in X.$$

A linear space (or a vector space) is thus a commutative group, whose elements, which we call vectors, can be multiplied by scalars (numbers). If the field of scalars is the real (complex) number field, we shall say that X is a real (complex) vector space.

Example 5.1. The set

$$\mathbb{R}^n := \{(x_1, x_2, \ldots, x_n)\,;\, x_1, x_2, \ldots, x_n \in \mathbb{R}\},$$

is a real linear space with the group operation and the scalar multiplication defined as

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

$$\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n), \qquad \lambda \in \mathbb{R}.$$

The group identity, $\mathbf{0}$, is of course the vector $\mathbf{0} := (0, 0, \ldots, 0)$, and the group inverse $-(x_1, x_2, \ldots, x_n) := (-x_1, -x_2, \ldots, -x_n)$.

In an analogous way, the set of complex n-tuples, $\mathbb{C}^n$, is a complex vector space under vector addition and scalar multiplication with complex numbers.

Exercise 5.1. Can the set $\mathbb{C}^n$ in a natural way be considered as a real vector space?

Example 5.2. The set of polynomials with real (complex) coefficients defined on the real line $\mathbb{R}$, $\mathcal{P}(\mathbb{R}, \mathbb{R})$ ($\mathcal{P}(\mathbb{R}, \mathbb{C})$), is a real (complex) vector space, if addition and scalar multiplication is defined in the natural way.

In the same way is the set of polynomials with real (complex) coefficients defined on the real line $\mathbb{R}$ of degree less then or equal to n, $\mathcal{P}_n(\mathbb{R}, \mathbb{R})$ ($\mathcal{P}_n(\mathbb{R}, \mathbb{C})$), a real (complex) vector space.

Example 5.3. Let I be an index-set and let $\mathcal{F}(I, \mathbb{R})$ ($\mathcal{F}(I, \mathbb{C})$) denote the set of real valued (complex valued) functions on I. Then $\mathcal{F}(I, \mathbb{R})$ ($\mathcal{F}(I, \mathbb{C})$) becomes a real (complex) vector space with pointwise addition and multiplication.

Let J be a finite index set. The sum

$$\sum_{k \in J} \lambda_k \mathbf{x}_k,$$

is called a **linear combination** of the set of vectors $\{\mathbf{x}_k\}_{k \in J}$. The elements of the set of scalars $\{\lambda_k\}_{k \in J}$ are called coefficients.

Definition 5.1. Let I be any index set (possibly non finite). A set $\{\mathbf{x}_k\}_{k \in I}$ of vectors in a vector space X, is **linearly independent** iff for any finite subset J of I and any linear combination

$$\sum_{k \in J} \lambda_k \mathbf{x}_k = \mathbf{0} \;\Rightarrow\; \lambda_k = 0 \text{ for all } k \in J.$$

If a set of vectors is not linearly independent, the set is called **linearly dependent**.

Note that, since we do not have a notion of distance in this setting, it does not make sense to add more than a finite number of vectors.

Exercise 5.2. Show that the set of homogeneous polynomials $\{x^k\}_{k\in\mathbb{N}}$ is a linearly independent set of vectors in the vector space of complex polynomials on the real line $\mathcal{P}(\mathbb{R},\mathbb{C})$.

Proposition 5.1. *A non empty set S of non zero vectors in a linear space X is linearly dependent iff (at least) one of them is a linear combination of some of the others.*

Proof. If the set of vectors is linearly dependent, then there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_N$, not all zero, and vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ from S such that

$$\sum_{k=1}^{N} \lambda_k \mathbf{x}_k = \mathbf{0}.$$

We may assume (by renumbering if necessary) that $\lambda_1 \neq 0$ and thus we get

$$\sum_{k=2}^{N} \frac{\lambda_k}{\lambda_1} \mathbf{x}_k = \mathbf{x}_1.$$

The other implication is immediate. □

Definition 5.2. If every vector in a vector space X can be written as a linear combination of elements from a subset S in X , we say that the subset S **span** X.

The subsets that are both linearly independent and span the whole space are particularly important.

Definition 5.3. A **basis** in a linear space X is a set, S, of linearly independent vectors that span X.

Example 5.4. The vectors $(1,0,0,\ldots,0)$, $(0,1,0,\ldots,0)$, $\ldots$, $(0,0,0,\ldots,1)$ constitute a basis in the vector space $\mathbb{R}^n$. It is called the natural or canonical basis of $\mathbb{R}^n$.

Example 5.5. The set of homogeneous polynomials $\{x^k\}_{k\in\mathbb{N}}$ is a basis in $\mathcal{P}(\mathbb{R},\mathbb{C})$.

It is a consequence of (and in fact equivalent to) the Axiom of Choice that *every vector space has a basis*. From the Axiom of Choice it also follows the important fact that given a linear space X and any linearly independent set of vectors $\{\mathbf{e}_i\}_{i\in J}$ in X, we can **complement** (i.e. possibly add more elements to it) this set to a basis.

Note that we can have many different sets of basis vectors for a given vector space.

Example 5.6. You might be familiar with the result that the set of vectors $\mathbf{a}^i := (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, n$ is a basis in $\mathbb{R}^n$ iff the determinant of the corresponding matrix $[a_{ij}]$ is non zero.

We call a vector space finite dimensional iff we can find a basis containing only a finite number of elements. Otherwise, we call it infinite dimensional.

The following so called **basis theorem** is fundamental.

Theorem 5.2. *The number of elements in any basis of a vector space X is the same as in any other basis for X, i.e. either X is infinite dimensional or any basis has the same finite number of elements.*

Remark. Actually one can prove that, also for infinite dimensional vector spaces, two different sets of basis vectors always have the same cardinality.

Proof. If the space is infinite dimensional every basis has an infinite number of elements. Thus we may assume that the space is finite dimensional.

Let $S_1 = \{\mathbf{e}_k\}_{k=1}^n$ be a set that span X (since X is finite dimensional there exists at least one such set), and let $S_2 = \{\mathbf{f}_k\}_{k=1}^m$, be a linearly independent set of vectors. We shall prove that then $m \leq n$.

We begin by noting that the set $\{\mathbf{f}_1, \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ span X and is linearly dependent. We conclude that we can eliminate one of the basis vectors coming from S_1, say $\mathbf{e}_1$ and still have a set that span X. Thus the set $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ will span X and be linearly dependent. Since the set S_2 is linearly independent, we conclude that we can eliminate one more vector from S_1 and still have a set that span X, and so on…

If we in this process were to run out of elements from S_1, we would reach a contradiction to the linear independence of S_2. Thus $m \leq n$. In particular this implies that any set of basis vectors has to be finite. If now S_1 and S_2 are two sets of basis vectors, by switching their roles, we get $n \leq m$, and thus $m = n$. □

Definition 5.4. Let X be a vector space. We define the **dimension** of the space, $\dim(X)$, as the number of elements in any basis for X.

Proposition 5.3. *Let X be a vector space, then a set of vectors $\{\mathbf{x}_i\}_{i\in I}$ is a basis iff every $\mathbf{x} \in X$ in a unique way, i.e. with a unique finite subset of basis vectors with corresponding unique coefficients, can be written as a linear combination of elements from $\{\mathbf{x}_i\}_{i\in I}$.*

Proof. The only thing that we have to prove is that the representation of a given element in the basis is unique. If we have two representations of a given element, we may, by setting some coefficients equal to zero, assume that they contain the same basis vectors and then the result follows from the linear independence. □

Definition 5.5. A non empty subset Y of a vector space X is called a **linear subspace** iff

$$\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 \in Y \quad \text{for all } \mathbf{x}_1, \mathbf{x}_2 \in Y \text{ and all scalars } \lambda_1, \lambda_2.$$

Definition 5.6. Let S be any non empty subset of a linear space X. The **linear hull** of S, $lh(S)$, is then defined as the set of all linear combinations of vectors from S.

Exercise 5.3. Show that for any non empty subset S of a linear space X, the linear hull, $lh(S)$, is a linear subspace in X.

When it is clear from the context that we are assuming a linear structure, we shall sometimes omit the prefix *linear* in linear subspaces, and simply talk about subspaces of linear spaces.

Exercise 5.4. Let X be a vector space and let M and N be two given subspaces. Show that the sum

$$M + N := \{m + n \in X\,;\ m \in M, n \in N\},$$

is a subspace of X.

Exercise 5.5. Let X be a vector space and let M and N be two given subspaces. Show that the intersection $M \cap N$ is a subspace of X.

Definition 5.7. If two given subspaces M and N, of a linear space X, are such that $M \cap N = \{\mathbf{0}\}$ and $M + N = X$, we call them (algebraically) **complementary subspaces**.

Theorem 5.4. *Let M and N be two linear subspaces of a given linear space X. Then M and N are complementary iff every vector $\mathbf{x} \in X$ in a unique way can be written as $\mathbf{m} + \mathbf{n} = \mathbf{x}$, where $\mathbf{m} \in M$ and $\mathbf{n} \in N$.*

Proof. If M and N are complementary it follows that every vector $\mathbf{x} \in X$ can be written as a sum $\mathbf{m} + \mathbf{n} = \mathbf{x}$, where $\mathbf{m} \in M$ and $\mathbf{n} \in N$. That this representation is unique follows from the fact that if $\mathbf{m} + \mathbf{n} = \mathbf{m}' + \mathbf{n}'$, then $\mathbf{m} - \mathbf{m}' = \mathbf{n}' - \mathbf{n}$ and thus both sides are zero due to the fact that $M \cap N = \{\mathbf{0}\}$.

On the other hand, the unique representation property immediately implies that the subspaces are complementary. □

Definition 5.8. If M and N are complementary subspaces in X we will write $X = M \oplus N$ and we say that X is the **direct sum** of M and N.

Theorem 5.5. *If M and N are complementary subspaces in a given vector space X, then* $\dim(M) + \dim(N) = \dim(X)$.

Proof. Let $\{\mathbf{m}_\alpha\}_{\alpha\in A}$ and $\{\mathbf{n}_\beta\}_{\beta\in B}$ be two sets of basis vectors in M and N respectively, then $\{\mathbf{m}_\alpha\}_{\alpha\in A} \cup \{\mathbf{n}_\beta\}_{\beta\in B}$ is a set of basis vectors for X. □

Exercise 5.6. Show that if $\{\mathbf{x}_i\}_{i=1}^m$ is a linearly independent set, then the subspace $lh\{\mathbf{x}_i\}_{i=1}^m$ has dimension m.

Definition 5.9. Let X and Y be two vector spaces over the same field, $\mathbb{F}$, of scalars. A function $T : X \to Y$ is called a **linear map** (or **vector space homomorphism**) iff

$$T(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2) = \lambda_1 T(\mathbf{x}_1) + \lambda_2 T(\mathbf{x}_2), \qquad \lambda_1, \lambda_2 \in \mathbb{F}, \quad \mathbf{x}_1, \mathbf{x}_2 \in X.$$

Note that the identity map on any linear space X, $I : X \ni \mathbf{x} \mapsto \mathbf{x} \in X$, is a linear map.

Example 5.7. A given $m \times n$ matrix A represents a linear map Φ_A from $\mathbb{R}^n$ to $\mathbb{R}^m$ by the definition $\Phi_A : \mathbb{R}^n \ni \mathbf{x} \mapsto A\mathbf{x} \in \mathbb{R}^m$.

Definition 5.10. We say that two vector spaces X and Y over the same field of scalars are (linearly) **isomorphic** iff we can find a bijective linear map $T : X \to Y$. If X and Y are isomorphic we shall write

$$X \simeq Y.$$

That the relation of being isomorphic is an equivalence relation on the set of vector spaces follows from the following exercises.

Exercise 5.7. Show that the inverse of a vector space isomorphism is a linear map.

Exercise 5.8. Show that the composition of two linear maps, when it is defined, is a linear map.

Exercise 5.9. Show that two isomorphic linear spaces are of the same dimension, i.e. they are either both infinite dimensional or they are of the same finite dimension.

Example 5.7 is more than an example. Let X and Y be linear spaces with dimensions $\dim(X) = n$ and $\dim(Y) = m$, and let $T : X \to Y$ be a linear map. If we select two sets of basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} \subset X$ and $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\} \subset Y$, we have corresponding isomorphisms

$$S_{\mathbf{e}} : X \ni x = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_n\mathbf{e}_n \mapsto (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n,$$

and

$$S_{\mathbf{f}} : Y \ni y = \beta_1\mathbf{f}_1 + \beta_2\mathbf{f}_2 + \cdots + \beta_m\mathbf{f}_m \mapsto (\beta_1, \beta_2, \ldots, \beta_m) \in \mathbb{R}^m.$$

The linear map T together with these specific choices of basis vectors in X and Y, thus induce a unique linear map $T_{\mathbf{e},\mathbf{f}} : \mathbb{R}^n \to \mathbb{R}^m$ such that the following diagram commutes.

$$\begin{array}{ccc} X & \xrightarrow{\;T\;} & Y \\ {\scriptstyle S_{\mathbf{e}}}\big\downarrow & & \big\downarrow{\scriptstyle S_{\mathbf{f}}} \\ \mathbb{R}^n & \xrightarrow[T_{\mathbf{e},\mathbf{f}}]{} & \mathbb{R}^m \end{array}$$

Where

$$T_{\mathbf{e},\mathbf{f}} : \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T \mapsto A\alpha,$$

and the matrix $A = [A_{ki}]$, using Einstein's summation convention, is defined by

$$T(\mathbf{e}_i) = A_{ki}\mathbf{f}_k, \qquad i = 1, 2, \ldots, n.$$

Connected with any linear map, $T : X \to Y$, are two natural subspaces.

Definition 5.11. The subset

$$\ker(T) := \{\mathbf{x} \in X\,;\; T(\mathbf{x}) = \mathbf{0}\},$$

is called the **kernel** of T.

Exercise 5.10. Show that the kernel of a linear map $T : X \to Y$ is a subspace of X.

The other natural subspace is of course the **image**

$$\mathrm{Im}(T) := \{T(\mathbf{x}) \in Y\,;\; \mathbf{x} \in X\}.$$

Exercise 5.11. Show that the image set, $\mathrm{Im}(T)$, of a linear map $T : X \to Y$ is a subspace of Y.

If Z is a subspace of a given vector space X, over the field of scalars F, we can define an equivalence relation on X by the rule

$$\mathbf{x}_1 \sim \mathbf{x}_2 \quad \text{iff} \quad \mathbf{x}_1 - \mathbf{x}_2 \in Z.$$

Exercise 5.12. Show that this indeed defines an equivalence relation.

We can furthermore define a natural vector space structure on the quotient set X/Z, coming from the vector space structure on X.

Exercise 5.13. Show that vector addition and scalar multiplication is well defined by

$$\lambda_1[\mathbf{x}_1] + \lambda_2[\mathbf{x}_2] := [\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2], \qquad \lambda_1, \lambda_2 \in F, \quad \mathbf{x}_1, \mathbf{x}_2 \in X,$$

on the set of equivalence classes $X/Z = \{[\mathbf{x}]\}_{\mathbf{x}\in X}$. Thus with this definition X/Z is a vector space over F.

Exercise 5.14. Show that if

$$Y \oplus Z = X,$$

then Y and X/Z are isomorphic.

Proposition 5.6. *Let X be a linear space and let Z be a subspace, then* $\dim(X/Z) + \dim(Z) = \dim(X)$.

Proof. We may assume that Z is finite dimensional.

Let $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k\}$ be a basis for Z. We complement it to a basis for X (that this can be done in general is a consequence of the Axiom of Choice), say $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k\} \cup \{\mathbf{y}_\alpha\}_{\alpha\in A}$.

Let $Y := lh\{\mathbf{y}_\alpha\}_{\alpha\in A}$. From the fact that $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k\} \cup \{\mathbf{y}_\alpha\}_{\alpha\in A}$ is a basis for X it follows that Z and Y are complementary subspaces in X.

The proposition now follows since Y and X/Z are isomorphic. □

The following is the **fundamental theorem for vector space homomorphisms** (or linear mappings).

Theorem 5.7. *Let X and Y be two given vector spaces over the same field of scalars F, and let $T : X \to Y$ be a vector space homomorphism. Then the quotient space $X/\ker(T)$ is isomorphic to the subspace $\mathrm{Im}(T)$ of Y, i.e.*

$$X/\ker(T) \simeq \mathrm{Im}(T).$$

Proof. Let $K = \ker(T)$ and set $\tilde{T}([\mathbf{x}]) := T(\mathbf{x})$. The function

$$\tilde{T} : X/K \to \mathrm{Im}(T)$$

is well defined since

$$T(\mathbf{x}_1) = T(\mathbf{x}_2) \Leftrightarrow T(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0} \Leftrightarrow \mathbf{x}_1 - \mathbf{x}_2 \in K \Leftrightarrow [\mathbf{x}_1] = [\mathbf{x}_2].$$

It is a vector space homomorphism since

$$\begin{aligned}\tilde{T}(\lambda_1[\mathbf{x}_1] + \lambda_2[\mathbf{x}_2]) &= T(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2) = \lambda_1 T(\mathbf{x}_1) + \lambda_2 T(\mathbf{x}_2)\\ &= \lambda_1 \tilde{T}([\mathbf{x}_1]) + \lambda_2 \tilde{T}([\mathbf{x}_2]).\end{aligned}$$

It is clearly surjective, and injectivity follows since

$$\tilde{T}([\mathbf{x}]) = \mathbf{0} \Leftrightarrow T(\mathbf{x}) = \mathbf{0} \Leftrightarrow \mathbf{x} \in K. \qquad \square$$

An immediate corollary to this theorem is the following result, known as the **dimension theorem**.

Theorem 5.8. *Let X and Y be two given vector spaces over the same field of scalars F, and let $T : X \to Y$ be a vector space homomorphism, then*

$$\dim(\ker(T)) + \dim(Y) = \dim(Y/\mathrm{Im}(T)) + \dim(X).$$

Proof. From Theorem 5.7, we conclude that $\dim(X/\ker(T)) = \dim(\mathrm{Im}(T))$, and thus Proposition 5.6 concludes the proof. $\square$

In particular we note that for finite dimensional spaces X and Y, the (finite) number $\dim(\ker(T)) - \dim(Y/\mathrm{Im}(T))$ is independent of the operator T. In fact, by the dimension theorem, it is equal to $\dim(X) - \dim(Y)$.

Definition 5.12. Let X and Y be two given vector spaces over the same field of scalars F, and let $T : X \to Y$ be a vector space homomorphism, such that either $\ker(T)$ or $Y/\mathrm{Im}(T)$ (or both) are finite dimensional. We then define the **index** of T, $\mathrm{Ind}(T)$, as

$$\mathrm{Ind}(T) := \dim(\ker(T)) - \dim(Y/\mathrm{Im}(T)).$$