Lecture Notes

for

MA 1140

Elementary Linear Algebra

by

Dipankar Ghosh

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Department of Mathematics
Indian Institute of Technology Hyderabad
Sangareddy, Kandi, Telangana - 502285, India

7th February to 11th March, 2019

Contents

Introduction

Linear algebra[1] is the branch of mathematics concerning linear equations such as

$$a_1x_1 + \cdots + a_nx_n = b$$

linear functions such as

$$(x_1, \ldots, x_n) \mapsto a_1x_1 + \ldots + a_nx_n$$

and their representations through matrices and vector spaces.

Linear algebra is central to almost all areas of mathematics. For instance, linear algebra is fundamental in modern presentations of geometry, including for describing basic objects such as lines, planes and rotations. Also, functional analysis may be basically viewed as the application of linear algebra to spaces of functions. Linear algebra is also used in most sciences and engineering areas, because it allows modeling many natural phenomena, and efficiently computing with such models. For nonlinear systems, which cannot be modeled with linear algebra, linear algebra is often used as a first-order approximation.

Until the 19th century, linear algebra was introduced through systems of linear equations and matrices. In modern mathematics, the presentation through vector spaces is generally preferred, since it is more synthetic, more general (not limited to the finite-dimensional case), and conceptually simpler, although more abstract.

[1] The introduction is noted from Wikipedia [2].

Chapter 1

Vector Spaces

MA 1140 is the study of ‘vector spaces’ and the ‘maps’ between them. For now, you may consider $\mathbb{R}^n$ as an example of a vector space, where $\mathbb{R}$ is the set of real numbers. Essentially, a vector space means a collection of objects, we call them vectors, where we can add two vectors, and what we get is a vector; we can multiply a vector by a scalar, and what we get is a vector; cf. Figure 2.

1.1 Vector Space

Definition 1.1. A set V of objects (called vectors) along with vector addition ‘+’ and scalar multiplication ‘·’ is said to be a **vector space** over a field $\mathbb{F}$ (say, $\mathbb{F} = \mathbb{R}$, the set of real numbers) if the following hold:

(1) V is closed under ‘+’, i.e. $x + y \in V$ for all $x, y \in V$.

(2) Addition is commutative, i.e. $x + y = y + x$ for all $x, y \in V$.

(3) Addition is associative, i.e. $(x + y) + z = x + (y + z)$ for all $x, y, z \in V$.

(4) Additive identity, i.e. there is $0 \in V$ such that $x + 0 = x$ for all $x \in V$.

(5) Additive inverse, i.e. for every $x \in V$, there is $-x \in V$ such that $x + (-x) = 0$.

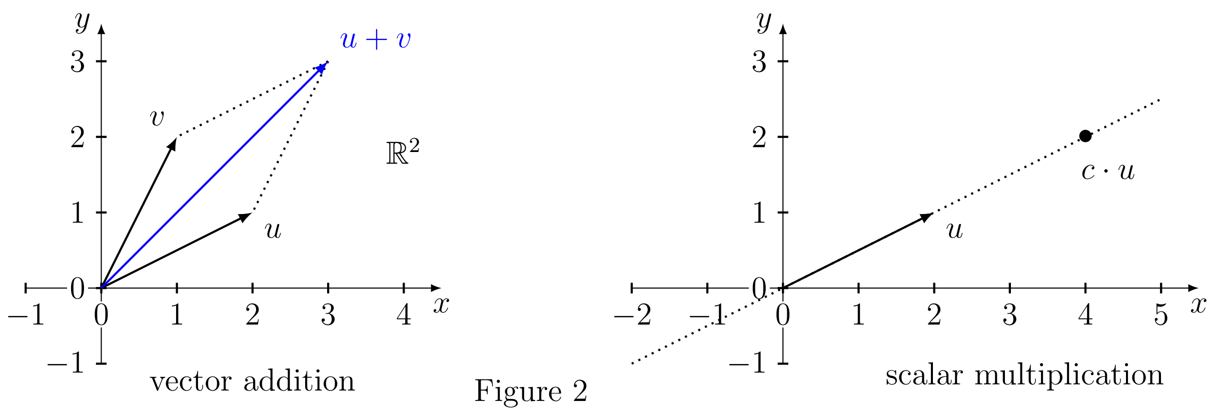


Figure 2

(6) V is closed under '$\cdot$', i.e. $c \cdot x \in V$ for all $c \in \mathbb{F}$ and $x \in V$.

(7) $1 \cdot x = x$ for all $x \in V$.

(8) $(ab) \cdot x = a \cdot (b \cdot x)$ for all $a, b \in \mathbb{F}$ and $x \in V$.

(9) $a \cdot (x + y) = a \cdot x + a \cdot y$ for all $a \in \mathbb{F}$ and $x, y \in V$.

(10) $(a + b) \cdot x = a \cdot x + b \cdot x$ for all $a, b \in \mathbb{F}$ and $x \in V$.

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

Remark 1.2. The first five properties are nothing but the properties of abelian group, i.e. $(V, +)$ is an abelian group.

We simply write cx instead of $c \cdot x$ for $c \in \mathbb{F}$ and $x \in V$ when there is no confusion.

From now, we work over the field $\mathbb{R}$.

Example 1.3. The following are examples of vector spaces.

(1) The n-tuple space, $V = \mathbb{R}^n$, where vector addition and scalar multiplication are defined as follows:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{and} \quad c \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}.$$

(2) The space $\mathbb{R}^{m \times n}$ of all $m \times n$ matrices

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} \quad \text{where } x_{ij} \in \mathbb{R}.$$

The vector addition and scalar multiplication are defined by component wise addition and multiplication as in (1).

(3) Let S be any non-empty set. Let V be the set of all functions from S into $\mathbb{R}$. The sum $f + g$ of two vectors f and g in V is defined to be

$$(f + g)(s) := f(s) + g(s) \text{ for all } s \in S.$$

The scalar multiplication $c \cdot f$ is defined by $(c \cdot f)(s) := cf(s)$. Clearly, V is a vector space. Note that the preceding examples are special cases of this one.

(4) The set $\mathbb{R}[x]$ of all polynomials $a_0 + a_1 x + \cdots + a_m x^m$, where $a_i \in \mathbb{R}$, x is an indeterminate and m varies over non-negative integers. The vector addition and scalar multiplication are defined in obvious way. Then $\mathbb{R}[x]$ is a vector space over $\mathbb{R}$.

Example 1.4. The set $V = \mathbb{R}^{n\times n}$ of all $n \times n$ matrices with vector addition defined by

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix} \times \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{pmatrix}$$
$$= \begin{pmatrix} -- & -- & -- \\ -- & (\sum_{k=1}^{n} x_{ik} y_{kj}) & -- \\ -- & -- & -- \end{pmatrix} \quad \text{(matrix multiplication)}$$

and scalar multiplication as before is NOT a vector space. Indeed, the operation '$\times$' is **not commutative** because

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Moreover, every matrix does not necessarily have multiplicative inverse. For example, there does not exist a matrix A such that

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \times A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Definition 1.5. A vector β in V is said to be a **linear combination** of vectors $\alpha_1, \alpha_2, \ldots, \alpha_r$ in V if $\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r$ for some $c_1, c_2, \ldots, c_r \in \mathbb{F}$.

Example 1.6. In $\mathbb{R}^2$, the vector $(1, 2)$ can be written as linear combinations of $\{(1, 0), (0, 2)\}$ and $\{(1, 1), (1, 0)\}$ respectively as

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} + (-1)\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

In view of Definition 1.5, one may ask the following questions. Suppose $\alpha_1, \alpha_2, \ldots, \alpha_r$ and β are given vectors in V. Is it possible to write β as a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_r$? If yes, then is it a unique way to write that? We find the answers to these questions as we proceed further.

Remark 1.7. In Figure 3 below, the set of all linear combinations of $\{u\}$ is given by the dotted line. So v cannot be written as a linear combination of $\{u\}$.

1.2 Subspaces

Definition 1.8. Let V be a vector space over a field $\mathbb{F}$. A subspace of V is a subset W of V which is itself a vector space over $\mathbb{F}$ with the same operations of vector addition and scalar multiplication on V.

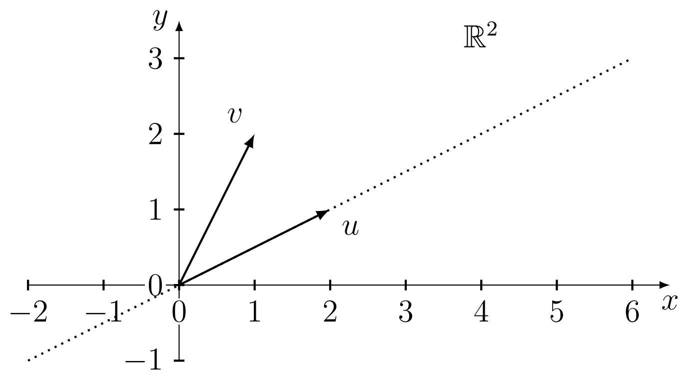


Figure 3

Theorem 1.9. Let W be a non-empty subset of a vector space V over $\mathbb{F}$. Then W is a subspace of V if and only if for each pair of vectors $\alpha, \beta \in W$ and each scalar $c \in \mathbb{F}$, the vector $c\alpha + \beta$ belongs to W.

Proof. Exercise! Note that many properties of W will be inherited from V. □

Theorem 1.10. Let V be a vector space over a field $\mathbb{F}$. The intersection of any collection of subspaces of V is a subspace of V.

Proof. Exercise! Use Theorem 1.9. □

Example 1.11. (1) The subset W consisting of the zero vector of V is a subspace of V.

(2) In $\mathbb{R}^n$, the set of n-tuples $(x_1, \dots, x_n)$ with $x_1 = 0$ is a subspace; while the set of n-tuples with $x_1 = 1$ is NOT a subspace.

(3) The set of all 'symmetric matrices' forms a subspace of the space of all $n \times n$ matrices. Recall that an $n \times n$ square matrix A is said to be symmetric if (i, j)th entry of A is same as its (j, i)th entry, i.e. $A_{ij} = A_{ji}$ for each i and j.

Definition 1.12. Let S be a set of vectors in a vector space V. The **subspace spanned** by S is defined to be THE smallest subspace of V containing S. We denote this subspace by $\mathrm{Span}(S)$.

What is the guarantee for existence of a subspace spanned by a given set? The following theorem is giving us that guarantee.

Theorem 1.13. Let S be a set of vectors in a vector space V. The following subspaces are equal.

(1) The intersection of all subspaces of V containing S.

(2) The set of all linear combinations of vectors in S, i.e. $\{c_1v_1 + \cdots + c_rv_r : c_i \in \mathbb{R},\ v_i \in S\}$. (One can check that it is a subspace.)

(3) The subspace $\mathrm{Span}(S)$, i.e. the smallest subspace of V containing S.

Proof. Let W_1, W_2 and W_3 be the subspaces described as in (1), (2) and (3) respectively. Then W_1 is contained in any subspace of V containing S. Therefore, since W_1 is a subspace (by Theorem 1.10), W_1 is the smallest subspace of V containing S, i.e. $W_1 = W_3$.

Using Theorem 1.9, it can be shown that W_2 is a subspace. Therefore, since W_2 contains S, we have $W_1 \subseteq W_2$. Notice that any subspace of V containing S also contains all linear combinations of vectors in S, i.e. any subspace of V containing S also contains W_2. Hence it follows that $W_2 \subseteq W_1$. Therefore $W_1 = W_2$. Thus $W_1 = W_2 = W_3$. □

Remark 1.14. In Figure 3, the subspace spanned by $\{u\}$ can be described by the dotted line; while the subspace spanned by $\{u, v\}$ is $\mathbb{R}^2$; see Example 1.19.

1.3 Basis and Dimension

Definition 1.15. Let V be a vector space over $\mathbb{R}$. A subset S of vectors in V is said to be **linearly dependent** if there exists (distinct) vectors $v_1, v_2, \ldots, v_r$ in S and scalars $c_1, c_2, \ldots, c_r$ in $\mathbb{R}$, not all of which are 0, such that

$$c_1 v_1 + \cdots + c_r v_r = 0.$$

A set S which is not linearly dependent is called **linearly independent**. If $S = \{v_1, v_2, \ldots, v_n\}$ is finite, we say that $v_1, v_2, \ldots, v_n$ are linearly dependent (or independent) instead of saying that S is so.

Remark 1.16. The following statements can be verified easily.

(1) Any set containing the 0 vector is linearly dependent.

(2) Every non-zero vector v in V is linearly independent.

(3) A set S of vectors is linearly independent if and only if every finite subset of S is linearly independent, i.e., if and only if for every $\{v_1, \ldots, v_r\} \subseteq S$,

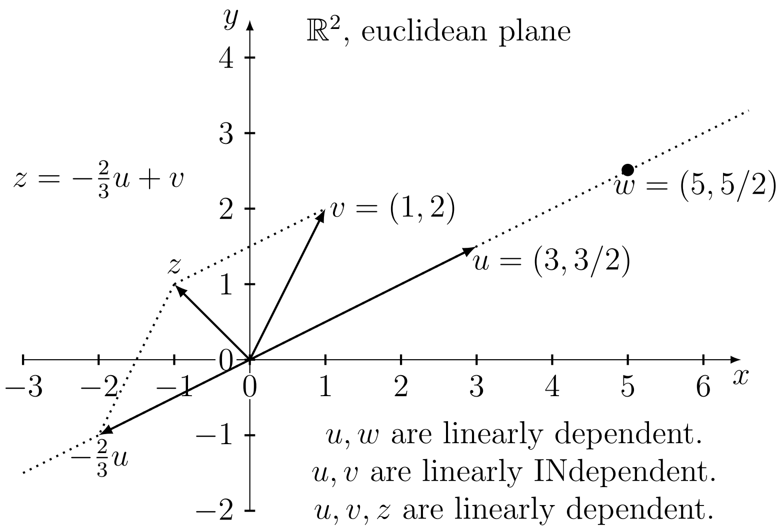
$$c_1 v_1 + \cdots + c_r v_r = 0 \quad \Longrightarrow \quad c_i = 0 \text{ for all } 1 \leqslant i \leqslant r.$$

(4) A finite set $\{v_1, \ldots, v_r\}$ of vectors is linearly independent if and only if $c_1 v_1 + \cdots + c_r v_r = 0$ implies that $c_i = 0$ for all $1 \leqslant i \leqslant r$.

(5) A set S of vectors is linearly dependent if and only if there exists at least one vector $v \in S$ which belongs to the subspace spanned by $S \setminus \{v\}$.

(6) Any set containing a linearly dependent subset is again linearly dependent.

(7) Any subset of a linearly independent set is linearly independent.



Definition 1.17. Let V be a vector space over $\mathbb{R}$. A set S of vectors in V is called a **basis** of V if S is linearly independent and it spans the space V (i.e., the subspace spanned by S is V). The space V is said to be finite dimensional if it has a finite basis.

Example 1.18. In $\mathbb{R}^2$, the vectors $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $v_3 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ are linearly dependent because $(7/3)v_1 + (1/3)v_2 + (-1)v_3 = 0$.

Example 1.19. The set $\left\{ v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}$ forms a basis of $\mathbb{R}^2$. Indeed, geometrically, it can be observed that v_1, v_2 are linearly independent, and $\{v_1, v_2\}$ spans $\mathbb{R}^2$. Or directly, we see that for EVERY vector $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$, the system

$$x \begin{pmatrix} 1 \\ 2 \end{pmatrix} + y \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{i.e.,} \quad \begin{cases} x + 2y = a \\ 2x + y = b \end{cases} \quad \text{or} \quad \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

has a UNIQUE solution in x, y because the coefficient matrix $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ is invertible. So every vector in $\mathbb{R}^2$ can be written as a linear combination of $\{v_1, v_2\}$, hence it spans the space $\mathbb{R}^2$. Moreover, when $a = b = 0$, then the system has THE trivial solution $x = y = 0$. Thus $\{v_1, v_2\}$ is linearly independent as well.

Example 1.20. In $\mathbb{R}^n$, let S be the subset consisting of the vectors:

$$\begin{aligned} e_1 &= (1, 0, 0 \ldots, 0) \\ e_2 &= (0, 1, 0, \ldots, 0) \\ &\vdots \\ e_n &= (0, 0, 0, \ldots, 1). \end{aligned}$$

Note that any vector $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$ can be written as a linear combination $x_1e_1 + \cdots + x_ne_n$. So S spans $\mathbb{R}^n$. Moreover, it can be shown that S is linearly independent. Therefore S is a basis of $\mathbb{R}^n$. This particular basis is called the **standard basis** of $\mathbb{R}^n$.

Exercise 1.21. Show that for $\mathbb{R}[x]$, the set of all polynomials over $\mathbb{R}$, the subset

$$S = \{x^n : n = 0, 1, 2, \ldots\}$$

forms a basis.

Lemma 1.22. Let V be a vector space over $\mathbb{R}$. Suppose $\{v_1, v_2, \ldots, v_n\}$ spans V. Let u be a non-zero vector in V. Then some v_i can be replaced by u to get another spanning set of V, i.e., if necessary, then after renaming the vectors $\{v_1, v_2, \ldots, v_n\}$, we obtain that $\{u, v_2, \ldots, v_n\}$ spans V.

Proof. Since $\{v_1, v_2, \ldots, v_n\}$ spans V, u can be written as

$$u = c_1v_1 + c_2v_2 + \cdots + c_nv_n \text{ for some } c_1, \ldots, c_n \in \mathbb{R}. \tag{1.1}$$

Hence, since $u \neq 0$, at least one $c_i \neq 0$. Therefore (1.1) yields that

$$v_i = (1/c_i)u + (c_1/c_i)v_1 + \cdots + (c_{i-1}/c_i)v_{i-1} + (c_{i+1}/c_i)v_{i+1} + \cdots + (c_n/c_i)v_n. \tag{1.2}$$

Since any vector v in V can be written as a linear combination of $v_1, v_2, \ldots, v_n$, using (1.2) in that linear combination, it follows that v can be written as a linear combination of $v_1, \ldots, v_{i-1}, u, v_{i+1}, \ldots, v_n$. Thus $\{v_1, \ldots, v_{i-1}, u, v_{i+1}, \ldots, v_n\}$ spans V. □

Theorem 1.23. Let V be a vector space over $\mathbb{R}$. Suppose V is spanned by a finite set $\{v_1, v_2, \ldots, v_n\}$ of n vectors, and $\{u_1, u_2, \ldots, u_m\}$ is a linearly independent set of m vectors in V. Then m is finite and $m \leqslant n$.

Proof. If possible, let $n < m$. Note that every $u_i \neq 0$. In view of Lemma 1.22, if necessary, by renaming the vectors $v_1, \ldots, v_n$, we have that $\{u_1, v_2, v_3, \ldots, v_n\}$ spans V.

In the 2nd step, since $u_2 \in V = \mathrm{Span}\{u_1, v_2, v_3, \ldots, v_n\}$, it follows that

$$u_2 = b_1u_1 + b_2v_2 + b_3v_3 + \cdots + b_nv_n \quad \text{for some } b_i \in \mathbb{R}.$$

Then at least one of $\{b_2, \ldots, b_n\}$ is non-zero, otherwise if $b_i = 0$ for all $2 \leqslant i \leqslant n$, then $\{u_1, u_2\}$ is linearly dependent, which contradicts the hypotheses. Thus, as in the proof of Lemma 1.22, if necessary, by renaming the vectors $v_2, \ldots, v_n$, we have that $\{u_1, u_2, v_3, \ldots, v_n\}$ spans V.

Continuing in this way, after n steps, we obtain that $\{u_1, u_2, \ldots, u_n\}$ spans V. Hence

$$u_{n+1} \in V = \mathrm{Span}\{u_1, u_2, \ldots, u_n\}.$$

This yields that $\{u_1, u_2, \ldots, u_{n+1}\}$ is linearly dependent, which contradicts the hypotheses. Therefore $m \leqslant n$. □

Corollary 1.24. If V is a finite dimensional vector space, then any two bases of V have the same number of elements.

Proof. Since V is finite dimensional, it has a finite basis $\{v_1, \ldots, v_n\}$. If $\{u_1, \ldots, u_m\}$ is a basis of V, then by Theorem 1.23, m is finite and $m \leqslant n$. By the same argument, $n \leqslant m$. Thus $m = n$. $\square$

Corollary 1.24 allows us to define the **dimension** of a finite dimensional vector space as the number of elements in a basis. We denote the dimension of a finite dimensional vector space V by $\dim(V)$. If V is not finite dimensional, then we set $\dim(V) := \infty$. Hence Theorem 1.23 can be reformulated as follows.

Corollary 1.25. Let V be a finite dimensional vector space and $d = \dim(V)$. Then

(1) any subset of V containing more than d vectors is linearly dependent.

(2) no subset of V containing fewer than d vectors can span V.

Exercise 1.26. Find a basis of the space $\mathbb{R}^{m\times n}$ of all $m \times n$ matrices, and hence $\dim(\mathbb{R}^{m\times n})$. Hint: In view of Example 1.20, we have $\dim(\mathbb{R}^n) = n$.

Lemma 1.27. Let S be a linearly independent subset of a vector space V. Suppose $v \notin \operatorname{Span}(S)$. Then $S \cup \{v\}$ is linearly independent.

Proof. Let $c_1v_1 + \cdots + c_rv_r + cv = 0$ for some (distinct) vectors $v_1, \ldots, v_r \in V$ and scalars $c_1, \ldots, c_r, c$. If $c \neq 0$, then it follows that $v \in \operatorname{Span}(S)$, which is a contradiction. Therefore $c = 0$, and hence $c_1v_1 + \cdots + c_rv_r = 0$. Since S is linearly independent, it follows that $c_i = 0$ for every $1 \leqslant i \leqslant r$. $\square$

Theorem 1.28. Let W be a subspace of a finite dimensional vector space V, and S be a linearly independent subset of W. Then S is finite, and it is part of a (finite) basis of W.

Proof. Note that S is also a linearly independent subset of V. So S contains at most $\dim(V)$ elements. If S spans W, then S itself is a basis of W, and we are done. If $\operatorname{Span}(S) \neq W$, then there is a vector $v_1 \in W \smallsetminus \operatorname{Span}(S)$. Hence, by Lemma 1.27, $S_1 := S \cup \{v_1\}$ is linearly independent. If $\operatorname{Span}(S_1) = W$, then we are done. Otherwise, there is $v_2 \in W \smallsetminus \operatorname{Span}(S_1)$, and hence by Lemma 1.27, $S_2 := S \cup \{v_1, v_2\}$ is linearly independent. This process stops after some finite steps because at most $\dim(V)$ linearly independent vectors can be there in W. So finally we obtain a set

$$S \cup \{v_1, v_2, \ldots, v_m\} \subset W$$

which spans W, i.e., it forms a basis of W. $\square$

As an immediate consequence, we obtain the following.

Corollary 1.29. In a finite dimensional vector space V, every linearly independent set of vectors is part of a basis.

Corollary 1.30. Let W be a PROPER subspace of a finite dimensional vector space V. Then W is finite dimensional and $\dim(W) < \dim(V)$.

Proof. Since W is a proper subspace of V, it follows that $V \neq 0$. So, if $W = 0$, then we are done. Thus we may assume that there is a non-zero vector w in W. Then, by Theorem 1.28, $\{w\}$ can be extended to a finite basis (say S) of W. So, in particular, W is finite dimensional. Since $\mathrm{Span}(S) = W \subsetneq V$, there is a vector $v \in V \smallsetminus \mathrm{Span}(S)$. By Lemma 1.27, $S \cup \{v\}$ is a linearly independent subset of V. Hence, again by Theorem 1.28, $S \cup \{v\}$ can be extended to a basis of V. Therefore $\dim(W) < \dim(V)$. $\square$

Definition 1.31. Let $W_1, W_2, \ldots, W_r$ be subspaces of a vector space V. Then the **sum of these subspaces** is defined to be

$$W_1 + W_2 + \cdots + W_r = \{w_1 + w_2 + \cdots + w_r : w_i \in W_i\}.$$

Remark 1.32. It can be verified that $W_1 + W_2 + \cdots + W_r$ is again a subspace of V. In fact, this is the subspace of V spanned by $W_1 \cup W_2 \cup \cdots \cup W_r$, i.e., the smallest subspace containing all $W_1, W_2, \ldots, W_r$.

Theorem 1.33. Let W_1 and W_2 are finite dimensional subspaces of a vector space V. Then $W_1 + W_2$

Bibliography

[1] K. Hoffman and R. Kunze, *Linear Algebra*, 2nd Edition.

[2] https://www.wikipedia.org

[3] G. Strang, *Linear Algebra and Its Applications*, 4th Edition.