

Here is a summary of concepts involved with vector spaces.

1. **Vector Space.** A **vector space** over a field k is a non empty set V together with a well defined addition ‘+’ and a scalar multiplication ‘ \cdot ’ by elements of the field k , satisfying the following axioms.

- **Operations.** For all $u, v \in V$ and $c \in k$ we have $u + v \in V$ and $cu \in V$. Sometimes, we write $c \cdot u$ but this \cdot is often omitted.
- **Addition properties.** The addition is commutative ($u + v = v + u$) and associative ($u + (v + w) = (u + v) + w$).
- **Additive identity and inverse.** There is a zero vector, denoted 0 such that $u + 0 = 0 + u = u$ for all u . Moreover, for each u there is a u^* such that $u + u^* = u^* + u = 0$. The u^* can be shown to be uniquely defined by u and is denoted as $-u$. The zero vector is also shown to be unique!
- **Distributivity and unitariness.** The two operations interact naturally:

$$c(u + v) = cu + cv, (c + d)u = cu + du, (cd)u = c(du).$$

Moreover, $1(u) = u$ for all u .

In this course, the field k is usually \mathfrak{R} the field of real numbers. In that case, we drop the phrase “over the field \mathfrak{R} ”.

2. **Subspace** If V is a vector space over k and W is a non empty subset, then we say that W is a **subspace** of V if we have:

- For all $w_1, w_2 \in W$ we have $w_1 + w_2 \in W$.
- For all $c \in k$ and $w \in W$ we have $cw \in W$.

We note that the vector 0 will always belong to a subspace as soon as it is non empty, since $w \in W$ implies $0w = 0 \in W$ by the second subspace condition above.

Hence, you may replace the condition of W being non empty by the simpler condition $0 \in W$, as done in the book.

3. **A challenging example.** Here is an exotic example of a vector space which should be studied to verify your understanding of the above definition.

Let $V = \mathfrak{R}$ be made into a vector space over the usual field of real numbers \mathfrak{R} as follows:

- We define a new addition \oplus on V by the formula:

$$v \oplus w = v + w - 1$$

where the operations on the right hand side are the usual operations in real numbers.

- We define a new scalar multiplication \odot by \mathfrak{R} on V by the formula:

$$c \odot v = cv + 1 - c$$

where, as before, the operations on the right are the usual operations in real numbers.

It is instructive to verify all the axioms from these definitions. You should also identify what $-v$ means. This example should be kept in mind while analyzing all the following concepts.

You should also make an enhanced version of this example by taking V to be \mathfrak{R}^n as the set, but using the above definitions of addition and scalar multiplication, suitably generalized. ¹

¹Can you make other examples of such weird operations? Here is a general hint and secret of all such constructions.

Let V be any vector space over a field k . Let ψ be any bijective map of V to itself.

Define a new vector space W which uses the same V as an underlying set but defines operations as follows.

$$w_1 \oplus w_2 = \psi^{-1}(\psi(w_1) + \psi(w_2)) \text{ and } c \odot w = \psi^{-1}(c\psi(w)).$$

It can be shown that W is a vector space “isomorphic” to V which means essentially the same as V . See below for explanation of “isomorphic”.

Can you guess the ψ for the displayed example?

4. **A Universal example.** Let S be any non empty set and consider

$$F_S^k = \{f : S \rightarrow k \mid \text{where } f(s) = 0 \text{ for all except finitely many } s \in S. \}$$

It can be shown that every vector space can be described in this manner, but finding such an explicit S can be tedious and it is better to use the basic definition.

If $k = \mathfrak{R}$ we may drop it from the notation.

It is easy to verify how F_S^k is a vector space by defining $(f + g)(s) = f(s) + g(s)$ and $cf(s) = cf(s)$. The extra condition on f is not necessary, but it is essential if you want to claim that every vector space has a standard structure!

5. **Basic examples.** Here are some of the standard examples.

- **Euclidean spaces.** The space k^n consisting of all n -tuples of elements of k , usually written as a column.

This can be described as F_S^k where S is the set $\{1, 2, 3, \dots, n\}$. A typical function f in the vector space may

be displayed as $\begin{pmatrix} f(1) \\ f(2) \\ \dots \\ f(n) \end{pmatrix}$. This leads to the usual notation for k^n .

- **The case of an infinite S.** If we take $S = \{1, 2, \dots, n, \dots\}$, the set of natural numbers, then we find it convenient to display $f \in F_S$ as

$$f(1) + f(2)x + f(3)x^2 + \dots + f(n+1)x^n + \dots$$

Note that the description of F_S implies that after some large enough exponent N , the coefficients are all zero and we have a set of polynomials.

The book denotes $F_S^{\mathfrak{R}}$ by the symbol \mathbb{P} . A general notation for F_S^k is also $k[x]$ which is the ring of polynomials in x with coefficients in k , where, we have chosen to ignore the usual multiplication of polynomials!

We now note that if we keep the same set $S = \{1, 2, \dots, n, \dots\}$ but drop the special condition on functions we get a much bigger set, namely

$$H = \{f : S \rightarrow k\}.$$

As before, any such $f \in H$ be displayed as $f(1) + f(2)x + f(3)x^2 + \dots + f(n+1)x^n + \dots$.

Since there is no special condition on the function, we now get power series! The general notation for this set is $k[[x]]$, the ring of power series in x with coefficients in k , where, as before, we ignore the usual product of power series.

It can be shown that $H = F_T^k$ for some convenient set T , but finding such a T is a daunting task!

There is also a well known subset of H when $k = \mathfrak{R}$, namely the set of convergent power series. To write it as $F_T^{\mathfrak{R}}$ is an even more daunting task!

6. **Basic structures in a vector space.** Now let V be a k -vector space (i.e. vector space over a field k).

For any subset $A \subset V$, we define its **span**:

$$\text{Span } A = \{c_1v_1, \dots + c_mv_m \mid \text{where } c_i \in k, v_i \in A \text{ and } m \text{ is some non negative integer.}\}.$$

Note that $\text{Span } A$ can be described as the set of all possible **linear combinations** of elements of A . Note that even when A is infinite, we only allow finitely many elements of it at a time! Also note that m is allowed to be zero and it gives the combination 0, by a standard convention.

We say that a set A **spans V or is a spanning set for V** if $\text{Span } A = V$.

A subset $A \subset V$ is said to be **linearly dependent** if there are elements $v_1, \dots, v_m \in A$ such that $c_1v_1 + \dots + c_mv_m = 0$ for some $c_1, \dots, c_m \in k$ with at least one non zero element c_i among them.

In application, other convenient forms of this condition are used. One such version is:

A subset $A \subset V$ is said to be linearly dependent if there is some $v \in A$ such that v is a linear combination of some elements $w_1, \dots, w_r \in A$ which are distinct from v . A compact way of saying this is to write that $v \in \text{Span } A \setminus \{v\}$.

A set $A \subset V$ is said to be **linearly independent**, if it is not linearly dependent.

We often drop the word “linearly” from these terms.

A subset $A \subset V$ is said to be a **basis** of V if

$$\text{Span } A = V \text{ and } A \text{ is linearly independent.}$$

We say that V is **finite dimensional** if it has a finite basis. The number of elements in a basis, is said to be **the dimension** of V . We write $\dim V$ or $\dim_k V$ if we wish to identify k .

We will soon argue that the dimension is a well defined number for any vector space, i.e. every basis of a vector space has the same number of elements.

However, for infinite dimensional spaces, we need to make a finer notion of “cardinality” which distinguishes between different infinite sets. ²

7. Homomorphisms, Isomorphisms and Automorphisms. Given k -vector spaces V, W a map $T : V \rightarrow W$ is said to be a **linear transformation** if it satisfies these two conditions:

- $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$.
- $T(cv) = cT(v)$ for all $c \in k$ and $v \in V$.

We may also use the term “homomorphism” (meaning similar structure) to denote such a map.

There are two concepts associated with the notion of a linear transformation (homomorphism).

First is “the image of T ” which can be formally denoted as

$$T(V) = \{T(v) \mid v \in V\}.$$

Second is “the Kernel of T ” which can be defined as:

$$\text{Ker } T = \{v \in V \mid T(v) = 0\}.$$

It is easy to verify that both the Kernel and the Image are respectively subspaces of V and W .

The homomorphism T is **injective** (or “**one to one**”) iff $\text{Ker } T = 0$ where we have used a slightly abused notation 0 in place of $\{0\}$. This abuse is routinely done!

The homomorphism T **surjective** (or “**onto**”) if $T(V) = W$ i.e. W is its total image.

The homomorphism T is said to be an **isomorphism** if it is both injective and surjective i.e. bijective. The word “iso” denotes sameness and the concept says that the two vector spaces with an isomorphism mapping one to the other are essentially the same. They, can be treated as replacements for each other in analyzing their properties.

An isomorphism of V to itself is called an automorphism.

8. A Fundamental Theorem. This is the fundamental theorem in the theory of vector spaces.

Let V be a vector space with a spanning set A . Then there is a subset B of A such that B is a basis of V .

We shall give a complete proof in case A is finite and a partial proof in case A is infinite.

Proof.

If A is independent, then it is a basis and we are done. If A is dependent, then there is some vector $v \in A$ which is a linear combination of vectors in $A_1 = A \setminus \{v\}$.

We now claim that $\text{Span } A = \text{Span } A_1$ by the following argument. **Proof of claim.** Note that any $w \in \text{Span } A$ can be written as: $w = cv + w_1$ where $c \in k$ and $w_1 \in \text{Span } A_1$.

By assumption, $v \in \text{Span } A_1$, so cv and hence w belongs to $\text{Span } A_1$. Thus $\text{Span } A \subset \text{Span } A_1$. Clearly $\text{Span } A_1 \subset \text{Span } A$ since $A_1 \subset A$.

This shows $\text{Span } A = \text{Span } A_1$.

²Two sets A, B are said to have the same cardinality if there is a bijective map from A to B . As the example of the sets $\{1, 2, 3, \dots, n, \dots\}$ and $\{2, 4, 6, \dots, 2n, \dots\}$ shows, a set can have the same cardinality as a proper subset. This suggests that one has to be very careful in dealing with infinite cardinalities.

Thus, if $V = \text{Span } A$ and A is dependent, then we get a proper subset A_1 of A such that $V = \text{Span } A_1$. We can now apply the same argument to A_1 and either get a basis or a smaller spanning set A_2 .

In case A is a finite set to begin with, this cannot continue indefinitely and we must get a basis at some stage.

We remark that we may run into an empty subset of A , in case the vector space is the zero space $\{0\}$. However, in this case the whole set A can only be $\{0\}$ or empty and we have nothing to argue!

It is also possible to do the above proof in a reversed process. We can start with independent subsets of A and enlarge them as much as possible. We argue that eventually we should get a basis.

In case the set A is infinite, we need a more general inductive principle called Zorn's Lemma. The proof proceeds by building a maximal subset A^* of A with the property that A^* is itself independent, but any set between A^* and A is dependent. Then it is easy to see that A^* is itself a basis for $\text{Span } A = V$.³

9. **Coordinates.** Let V be a vector space with a basis A . We show that V is isomorphic to F_A^k .

Note that by the definition of the basis, every vector $w \in V$ has a unique expression $w = c_1v_1 + \dots + c_rv_r$ for some $v_1, \dots, v_r \in A$ and $c_1, \dots, c_r \in k$.

For any $v \in A$, we define the v -coordinate of w with respect to A by the formula:

$$w_{[v,A]} = c_i \text{ if } v = v_i \text{ and } 0 \text{ if } v \text{ is not among the } v_1, \dots, v_r.$$

Now we define the map $\Phi : V \rightarrow F_A^k$ by

$$\Phi(w)(v) = w_{[v,A]}.$$

It is not hard to see that this is indeed an isomorphism!

The main trick is to note that from the definition of F_A^k , given any $f \in F_A^k$ we can see that the element $w = \sum_{v \in A} f(v)v$ is a well defined member of our vector space $\text{Span } A$. Our definition of Φ makes $\Phi(w) = f$. This shows surjectivity.

Injectivity is obvious.

Thus, we have shown why every vector space is essentially of the form F_S^k for some set S .

10. **Dimension formulas.** The next important result in vector spaces is that any basis of a vector space has exactly the same number of elements (interpreted as the same cardinality for infinite bases).

This number is the dimension of the vector space.

We only describe the argument when we have one finite basis.

Let A, B be two bases of a vector space V where A consists of n vectors. We shall construct a sequence of bases $A_0, A_1, A_2, \dots, A_n$ such that $A_0 = A$ but $A_n \subset B$ and each A_i has exactly n of elements. Now, since A_n is a basis and contained in the basis B , it must be equal to B , thus B has n elements as well.

We shall arrange our A_i to have **at least** their first i elements coming from B and we shall prove this by induction on i .

Thus the case $i = 0$ is clear, since $A_0 = A$ satisfies the condition.

Assume that $0 < m < n$ and we have A_m with its first m elements in B . Let V_1 be the span of the first m elements of A_m . Clearly V_1 is strictly smaller than V , since otherwise the basis A_m would be contained in V_1 making it dependent.

Now, by the same reasoning, we must have at least one vector $w \in B$ which is not in V_1 and let us write it as a combination of elements of A_m , say $w = u + v$ where u is a combination of the first m elements of A_m and v of the remaining elements of A_m . Now, $v \neq 0$ for otherwise the independence of B is contradicted.

Consider $A_{m+1} = A_m \cup \{w\} \setminus \{v_i\}$ where v_i is one of the vectors appearing in the combination of v . It is easy to see that A_{m+1} is a new basis with $m + 1$ elements from B , namely w and the first m elements of A_m . We shall make them the first $m + 1$ elements of A_{m+1} .

We are thus finished by induction!

Now we come to the most useful theorem in dimension theory of vector spaces.

³The idea of this proof is as follows. Suppose we have a subset A^* of A which is independent but which is maximal, in the sense any bigger subset of A containing A^* is dependent. In this case, it is not hard to see that $\text{Span } A^* = \text{Span } A$. Thus A^* is the desired basis.

Theorem. Let $T : V \rightarrow W$ be a surjective linear transformation of vector spaces.

Then

$$\dim V = \dim W + \dim \text{Ker } T.$$

Proof. Select a basis A_1 of $\text{Ker } T$ and a set A_2 such that $T(A_2)$ gives a basis of W . (This is done using the surjectivity of T .)

Now it is possible to argue that $A = A_1 \cup A_2$ is a basis of V . Here is the outline.

First show that $\text{Span } A = V$. Let $v \in V$. Then $T(v) \in \text{Span } T(A_2)$. Thus we have $u_2 \in \text{Span } A_2$ such that $T(v) = T(u_2)$ or $v - u_2 = u_1 \in \text{Ker } T$.

Then $u_1 \in \text{Span } A_1$ and thus $v = u_1 + u_2$ is in $\text{Span } A_1 \cup A_2$.

The independence of the set A is argued thus. Suppose some non trivial combination of elements of A is zero. Write the combination as $u_1 + u_2 = 0$ where $u_1 \in \text{Span } A_1$ and $u_2 \in \text{Span } A_2$.

Taking the image by T we see that

$$T(u_1 + u_2) = T(u_1) + T(u_2) = 0 + T(u_2).$$

Since $T(A_2)$ is a basis of W , we see that $T(u_2)$ is a trivial combination and hence u_2 is also a trivial combination.

Thus $u_2 = 0$.

From independence of A_1 it follows that u_1 is also the trivial combination.

Now note that A_1, A_2 are disjoint sets, since otherwise a basis of W will contain a zero vector (the image of a common element).

Hence, for finite dimensional V we get that the number of elements of A equals the sum of the number of elements of A_1 and A_2 .

This proves the formula.

Note: Our proof did not use the finite dimensionality of the vector spaces. It is needed only to count the elements of A as those of A_1 and those of A_2 .

11. Connection with \mathbf{k}^n .

Consider a finite dimensional k -vector space V and a basis B which we write as a formal vector

$$B = (v_1 \quad \cdots \quad v_n)$$

of vectors in V .

We construct a map $V \rightarrow k^n$ as follows. Given $v = c_1v_1 + \cdots + c_nv_n \in V$ we shall write

$$[v]_B = \begin{pmatrix} v_1 \\ \cdots \\ v_n \end{pmatrix} \in \mathfrak{K}^n.$$

The resulting vector shall be called the coordinate vector of v **with respect to the ordered basis** B .

This is best remembered by a natural identity:

$$v = B[v]_B.$$

This is to be interpreted as the usual product of a row with a column, generalized for our vector entries.

If $C = (w_1 \quad \cdots \quad w_n)$ is also a basis of V , then we get the identities:

$$w_1 = B[w_1]_B, \cdots, w_n = B[w_n]_B$$

and putting these together, we may write:

$$C = (w_1 \quad \cdots \quad w_n) = B ([w_1]_B \quad \cdots \quad [w_n]_B).$$

The matrix $([w_1]_B \ \cdots \ [w_n]_B)$ is a transformation matrix between bases B and C and may be conveniently denoted as P_B^C . Thus we have a suggestive notational display:

$$C = BP_B^C.$$

It can be imagined that the B at the foot of P_B^C cancels B and leaves us C .

Now we use this to compare the B and C coordinates for the same vector in $v \in V$.

Note:

$$B[v]_B = v = C[v]_C = BP_B^C[v]_C.$$

Comparing the first and the last terms we see:

$$[v]_B = P_B^C[v]_C.$$

12. Matrix of a transformation.

Finally, we make the matrix for a transformation $T : V \rightarrow W$ where we have chosen specific bases B for V and C for W .

We can write $T(B) = CM$ for some matrix M which is constructed as follows. Let the vectors in B be v_1, v_2, \dots, v_m in order and write $w_1 = T(v_1), \dots, w_m = T(v_m)$.

Then the i -th column of M is simply the C -coordinate vector of w_i , i.e. $[w_i]_C = [T(v_i)]_C$. Thus we see that

$$T(B) = (T(v_1) \ \cdots \ T(v_m)) = CM.$$

For $v \in V$ we see that $v = B[v]_B$ and hence $T(v) = T(B)[v]_B$ so that:

$$T(v) = T(B)[v]_B = CM[v]_B$$

and this shows that

$$[T(v)]_C = M[v]_B.$$

This shows that on the coordinate level, the map T is simply a multiplication by M . This imitates the work in \mathfrak{R}^n .