Vector spaces

A. Eremenko

August 23, 2024

First of all, we have to discuss numbers. Numbers in Linear Algebra must be elements of some *field*. A field is a set of objects with four arithmetic operations defined on them (addition, subtraction, multiplication and division). These operations must satisfy the usual properties (commutativity, associativity, distributivity). A field must contain two *distinct* elements 0 and 1 with the usual properties $x + 0 = x$ and $1x = x$ for all elements x of the field. I do not give a formal list of all those axioms, since in this course we use only two fields: the field of real numbers $\mathbf{R}$ and the field of complex numbers $\mathbf{C}$. The field $\mathbf{R}$ must be familiar to you, while complex numbers will be discussed in details later.

Here we give few other examples of fields. All of them can be used as sets of numbers for Linear Algebra.

1. Rational numbers $\mathbf{Q}$ form a field. Integers do not make a field, because for example 1/2 is is not an integer, so one cannot divide 1 by 2 in integers.

2. The simplest field of all is the field of two elements. It consists of 0 and 1. Addition and multiplication are defined as follows:

$$0 + 1 = 1 + 0 = 1,\ 0 + 0 = 1 + 1 = 0,\ 0 \cdot 1 = 1 \cdot 0 = 0,\ 0 \cdot 0 = 0,\ 1 \cdot 1 = 1.$$

You can check that so defined operations satisfy all usual rules of arithmetic. So one can do Linear Algebra dealing with vectors and matrices whose elements are 0 and 1 only.

But we restrict our attention to the fields of real and complex numbers only.

Vector spaces. For a given number field k, a *vector space* V *over* k is a set on which two operations are defined, subject to the properties listed below. The elements of V are called *vectors*. One operation is addition of vectors, another is multiplication of a vector by a number. In multiplication, we usually write the number on the left of the vector. The axioms are the following: for every vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and numbers c, c_j we have

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a},$$

$$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c},$$

There is a vector $\mathbf{0}$ such that

$$\mathbf{a} + \mathbf{0} = \mathbf{a} \quad \text{for every vector} \quad \mathbf{a}.$$

$$(c_1 + c_2)\mathbf{a} = c_1\mathbf{a} + c_2\mathbf{a},$$

$$c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}.$$

$$(c_1c_2)\mathbf{a} = c_1(c_2\mathbf{a}).$$

$$1 \cdot \mathbf{a} = \mathbf{a}.$$

$$0 \cdot \mathbf{a} = \mathbf{0}.$$

In the last equation, in the LHS we have the number 0 while in the RHS we have the zero-vector, whose existence is postulated in the 3-d axiom.

All results of Linear Algebra can be derived from these axioms. The definition of a vector space depends on the number field, so we distinguish real and complex vector spaces, depending on what number system is used.

The principal examples of vector spaces in this course are the spaces of matrices $\mathrm{Mat}_k(m \times n)$ and the spaces of functions defined on some set X with values in the field k. Operations on these spaces are defined in the usual way, and one can easily verify that they have the above-stated properties. Matrices with real entries form a real vector space. Matrices with complex entries can be treated as both a real and a complex vector space. In the first case only multiplication of them by real numbers is allowed. Similarly with functions: they can take real or complex values. Complex valued functions can be multiplied on either complex or real numbers (so they form two vector spaces, one real another complex). Real functions can be multiplied on real numbers only.

Important special cases of matrix spaces are $\mathbf{R}^n$, the real vector space of columns of n real numbers. Similarly, $\mathbf{C}^n$ is the space consisting of columns of complex numbers.

Few more comments on spaces of functions. Functions on a set X are added pointvise: the sum $f + g$ of two functions f and g defined on the same set X is a function which to every $x \in X$ puts into correspondence the number $f(x) + g(x)$. Similarly, if c is a number then cf is a function which to every $x \in X$ puts into correspondence the number $cf(x)$. The zero vector in the space of functions is the function which assigns the value 0 to every $x \in X$. (They also say that this function is *identically* equal to zero).

The simplest vector space of all is the space which consists of the single vector $\mathbf{0}$. The operations are of course defined as $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $c\mathbf{0} = \mathbf{0}$ for every number c. This space is called *trivial*.

An important method of defining new vector spaces from old ones is considering *subspaces*. A subspace U of a vector space V is a subset $U \subset V$ which is closed with respect to both operations. "Closed" means that:

a) for every $\mathbf{x} \in U$ and every $\mathbf{y} \in U$ the sum $\mathbf{x} + \mathbf{y}$ also belongs to U, and

b) for every $\mathbf{x} \in U$ and every number c, the vector $c\mathbf{x}$ also belongs to U.

In other words, a subset $U \subset V$ is a subspace if U itself is a vector space with the same operations defined in V.

Examples. We use the notation:

$$\{\mathbf{x} \in V : \text{ condition } \}$$

for "the set of all $\mathbf{x}$ in V which satisfy certain condition".

0. Every vector space V has the trivial space as a subspace. V itself is also a subspace.

1. The set of all positive numbers is not a subspace of $\mathbf{R}^1$. (Because when we multiply a positive number by the number -1 we do not obtain positive number, so property b) is violated).

2. A line through the origin in the plane $\mathbf{R}^2$ is a subspace since it satisfies a) and b) above: the sum of two vectors in this line also belongs to this line, and the product of a vector in this line also belongs to it.

3. A line which does not pass through the origin in $\mathbf{R}^2$ is not a subspace (since it does not contain the origin; the zero vector of the space must belong to every subspace).

4. The same as in examples 2,3 is true in $\mathbf{R}^3$.

5. The set of upper triangular $n \times n$ matrices is a subspace of the space $\mathrm{Mat}(n \times n)$. Same about lower triangular matrices.

6. The set of symmetric matrices is a subspace of $\mathrm{Mat}(n \times n)$. Same about skew-symmetric matrices. But the set of orthogonal matrices is not a subspace, since, for example $I - I = \mathbf{0}$; I is orthogonal, but $\mathbf{0}$ is not.
7. The set of all column vectors whose first coordinate is zero is a subspace of $\mathbf{R}^n$. The set of all column vectors whose first coordinate is 1 is not a subspace (again this set does not contain the zero vector).

8. The set of all continuous functions on an interval (a, b) is a subspace of the space of all functions on this interval. Indeed, it is proved in calculus that sum of two continuous functions is continuous, and a continuous function multiplied by a number is also continuous.

9. Same about the sets of all differentiable functions on an interval (a, b), or twice differentiable etc.

10. The set of polynomials is a subspace of the space of all functions defined on the real line (or on any interval). Polynomials will frequently occur in this course. Polynomials are functions of the form

$$f(x) = a_0 + a_1 x + \ldots + a_d x^d,$$

where a_j are numbers. If $a_d \neq 0$ we say that the polynomial has degree d.

9. The set of all polynomials of degree d is not a subspace of the space of all polynomials. (Sum of two polynomials of degree d is not necessarily of degree d).

10. But the set of all polynomials of degree **at most** d is a subspace of the space of all polynomials (and of the space of all functions on the real line).

Span, linear dependence, basis and dimension. Let V be a vector space, and $\mathbf{a}_1, \mathbf{a}_1, \ldots, \mathbf{a}_k$ some vectors in it, and $c_1, \ldots, c_k$ are numbers. The vector

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \ldots + c_k\mathbf{a}_k$$

is called a *linear combination* of $\mathbf{a}_1, \ldots, \mathbf{a}_k$. The set of all linear combinations of a set of vectors is called the *span* of this set of vectors.

To determine whether some vector $\mathbf{b}$ belongs to the span of $\mathbf{a}_1, \ldots, \mathbf{a}_n$ one has to find out whether there exist constants $c_1, \ldots, c_n$ such that

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \ldots + c_k\mathbf{a}_k = \mathbf{b}.$$

If $V = \mathbf{R}^n$, then $\mathbf{a}_j$ and $\mathbf{b}$ are column vectors, and the question is equivalent to the question whether the system

$$A\mathbf{x} = \mathbf{b}$$

has a solution. We know how to answer this question.

As a trivial example, the vectors $e_1, \ldots, e_n$ (the columns of the unit matrix) span $\mathbf{R}^n$. Indeed every vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbf{R}^n$ can be written as the linear combination

$$\mathbf{x} = x_1 e_2 + \ldots + x_n e_n.$$

Another example: vectors $1, x, x^2, \ldots$ in the space of all polynomials span this space. Notice that if you remove one vector from this set, the remaining set will not span the space of all polynomials.

Span of any system of vectors is a subspace. Check this. If a system of vectors spans the whole space, then any larger system (any system that contains our system) also spans the space.

If some set of vectors span the space, and one of them is a linear combination of others, then those others also span the space.

Notice! A linear combination is always a finite sum (we have not defined infinite sums!). But when we are talking about systems that span something these systems can be infinite, like in the examples with polynomials above.

A system of vectors is called *linearly dependent*, if there exist vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in this system and constants $c_1, \ldots, c_n$ *not all equal to zero* such that

$$c_1\mathbf{a}_1 + \ldots + c_n\mathbf{a}_n = \mathbf{0}.$$

If a set of vectors is not linearly dependent it is called *linearly independent*. In other words, a set is linearly independent if no linear combination of this set, except the trivial one, is equal to the zero vector.

Any system containing the zero vector is linearly dependent. Indeed, you can multiply the zero vector on 1 and the rest on zeros and obtain a linear dependence.

If a set of vectors is linearly dependent, then at least one of these vectors is a linear combination of the rest.

If a system of vectors is linearly dependent, then after adding some vectors to it it remains linearly dependent.

Examples. 1. If $\mathbf{a}_j$ are column vectors in $\mathbf{R}^n$, then linear dependence is equivalent to the system $A\mathbf{x} = \mathbf{0}$ to have a non-trivial solution. "Non-trivial" means "not the zero vector". Here $A = [\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_n]$, the matrix whose columns are $\mathbf{a}_j$. We already know how to decide whether such a system has a non-trivial solution: for existence of non-trivial solution it is necessary and sufficient that there are free variables in the REF of A. Since this is important, I recall several criteria:

a) If $A \in \mathrm{Mat}(m \times n)$ and $n > m$ (the number of variables is greater than the number of equations, then the system $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution.

b) If A is a square matrix, then $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution if and only if A is singular (=non-invertible).

2. Monomials $1, x, x^2, \ldots$ are linearly independent. To show this we argue by contradiction. Suppose they are linearly dependent. This means that some linear combination

$$c_0 + c_1 x + \ldots + c_d x^d$$

with not all $c_j = 0$ is equal to 0 identically. Plugging $x = 0$ we obtain that $c_0 = 0$. Then differentiate this identity and lug $x = 0$. We obtain that $a_1 = 0$. Continuing this argument we prove that all $c_j = 0$, contradiction.

A set of vectors in a space V is called a *basis* of the space V if this set is

a) linearly independent, and

b) spans the space V

For example, vectors $e_1, \ldots, e_n$ form a basis of $\mathbf{R}^n$.
Monomials $1, x, x^2, \ldots$ form a basis of the space of all polynomials.
Monomials $1, x, x^2, \ldots, x^d$ form a basis of the space of all polynomials of degree at most d.

Definition. *A vector space V is called finite-dimensional, if it is spanned by a finite system of vectors.*

For example, $\mathbf{R}^n$ is finite-dimensional, while the space of all polynomials is infinite-dimensional.

Theorem 1. *Every vector space has a basis.*

Here we use the convention that the basis of the trivial space is the empty set, so the theorem has no exceptions.

The full proof of this theorem is beyond the scope of this course. We will prove it only for finite-dimensional spaces.

Proof for finite-dimensional spaces. By definition, the space is spanned by a finite set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. If they are linearly independent then they form a basis. If they are linearly dependent, then some of them, for example $\mathbf{a}_1$ is a linear combination of the rest. But then $\mathbf{a}_1$ can be removed and the remaining vectors still span the space. Continuing this argument we eventually obtain a basis.

This theorem actually gives a way to construct a basis, once we know a finite spanning set.

Theorem 2. *In a finite-dimensional space V, all bases have the same number of vectors.*

This number is called the dimension of V, and denoted by $\dim V$.

Proof. Proving by contradiction, assume that we have two bases, $\mathbf{u}_1, \ldots, \mathbf{u}_m$ and $\mathbf{v}_2, \ldots, \mathbf{v}_n$. Then, by definition of a basis, each $\mathbf{v}_j$ is a linear combination of $\mathbf{u}_1, \ldots, \mathbf{u}_m$:

$$\mathbf{v}_j = \sum_{i=1}^{m} a_{i,j} \mathbf{u}_i, \quad j = 1 \ldots, n$$

We will show that $\mathbf{v}_j$ must be linearly dependent. Consider a linear combination of $\mathbf{v}_j$ with coefficients x_j:

$$\sum_{j=1}^{n} x_j \mathbf{v}_j = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{i,j} \mathbf{u}_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{i,j} x_j \right) \mathbf{u}_i.$$

We want this to be zero, so set all expressions in parentheses to be zero:

$$\sum_{j=1}^{n} a_{i,j} x_j = 0, \quad 1 \leq i \leq m$$

This is a system of $m < n$ homogeneous equations with n variables. So it has a non-trivial solution. Thus we obtained a linear dependence between $\mathbf{v}_j$.

To find dimension of a space, one has to find a basis.

Examples.

1. $\dim \mathbf{R}^n = n$. A basis is $e_1, \dots, e_n$.

2. $\dim \mathrm{Mat}(m \times n) = mn$. A basis consists of matrices whose one entry is 1 and the rest are zeros. There mn such matrices.

3. Dimension of the space of polynomials of degree at most d is $d+1$. A basis is formed by monomials.

4. Dimension of the space of symmetric matrices of size $n \times n$ is $n(n+1)/2$. For example, when $n = 2$ a basis is

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

5. Dimension of the space of skew-symmetric matrices of size $n \times n$ is $n(n-1)/2$.

The main use of a basis is the possibility to represent every vector in the space by its *coordinates*. Let V be a vector space of dimension n, and $\mathbf{v}_1, \dots, \mathbf{v}_n$ a basis. Then every vector $\mathbf{x} \in V$ can be written as a linear combination

$$\mathbf{x} = \sum_{j=1}^{n} c_j \mathbf{v}_j.$$

If $(\mathbf{x} = (x_1, \dots, x_n)^T$ and we choose the standard basis in $\mathbf{R}^n$, then this becomes

$$\mathbf{x} = x_1 e_2 + \ldots + x_n e_n.$$

So x_j are coordinates of $\mathbf{x}$ in the standard basis.
Claim. *This representation is unique.*

Proof. Suppose we have two such distinct representations of the same vector:

$$\sum_{j=1}^{n} c_j \mathbf{v}_j = \sum_{j=1}^{n} c'_j \mathbf{v}_j.$$

Then

$$\sum_{j=1}^{n} (c_j - c'_j) \mathbf{v}_j = \mathbf{0}$$

And we must have $c_j = c'_j$ for every j, since $\mathbf{v}_j$ are linearly independent.

Coefficients c_j are called the *co-ordinates* of the vector **c** with respect to the basis $\mathbf{v}_1, \ldots, v_n$.

To expand a vector **b** in terms of a basis $v_1, \ldots, v_n$, we have to solve the system $A\mathbf{c} = \mathbf{b}$, where $A = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ the matrix whose columns are basis vectors (compare Lecture 1).

Every non-trivial vector space has infinitely many bases. In physics, a basis is called a "coordinate system", or a "frame of reference". It allows to describe every vector in the space by numbers (its coordinates).

Change of the basis. Suppose we have two bases $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Then every vector **x** can be expanded in either basis:

$$\mathbf{x} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = b_1\mathbf{u}_1 + \ldots + b_n\mathbf{u}_n.$$

What is the relation between the coordinates c_j and b_j?

To answer this question we expand all vectors of the u-basis in terms of the v-basis:

$$\mathbf{u}_j = \sum_{i=1}^{n} a_{i,j}\mathbf{v}_i, \quad 1 \leq j \leq n. \tag{1}$$

Notice how we write this expansion: summation is with respect to the *first* subscript! We obtain a square matrix $A = (a_{i,j})$ which is called the matrix of the u-basis in terms of the v-basis, or the transition matrix. Now we substitute expressions (1) into the u-expansion of **x**:

$$\mathbf{x} = \sum_{j=1}^{n} b_j\mathbf{u}_j = \sum_{j=1}^{n} b_j \sum_{i=1}^{n} a_{i,j}\mathbf{v}_i,$$

and regroup to obtain the coefficients at $\mathbf{v}_i$:

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{i,j} b_j \right) \mathbf{v}_i.$$

The expressions in parentheses are coordinates of **x** in the v-basis.

Thus we have

$$\mathbf{c} = A\mathbf{b}, \tag{2}$$

where **c** is the column of coefficients of **x** in v-basis, **b** is the column of the coefficients of **x** in the u-basis, and A is the matrix of the u-basis in terms of the v basis.

This is called the change of the basis formula.

Theorem 3. *Let U be a subspace of a finite-dimensional space V. Then any basis of U can be extended to a basis of V.*

Proof. Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be a basis in U. If it spans V then it is a basis in V and we are done. If it does not span V, then there is a vector $\mathbf{v}_{k+1}$ in V which is not a linear combination of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Then $\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}$ must be linearly independent. Indeed, if there is a linear dependence between them, then the coefficient at $\mathbf{v}_{k+1}$ cannot be zero (since $v_1, ldots, v_k$ are linearly independent, nd then $\mathbf{v}_{k+1}$ will be a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Continuing this procedure of adding vectors, we eventually obtain a linearly independent system which spans V and begins with $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Here are few applications of the concept of dimension.

Theorem 4. *If U is a subspace of V and they have the same finite dimension, then $U = V$.*

Theorem 5. *If $\dim V = n$ and $v_1, \ldots, v_n$ is a linearly independent set, then this set is a basis. If $u_1, \ldots, u_n$ span V then they form a basis.*

So when we know the dimension of the space only one condition is sufficient to verify to make sure that a set of vectors forms a basis.