Vector Spaces

Daoping Zhang

The Chinese University of Hong Kong

dpzhang@math.cuhk.edu.hk

April 16, 2020

Outline

1. Definitions
2. Linear Combinations
3. Convex Sets
4. Linear Independence
5. Dimension
6. The Rank of a Matrix

A collection of objects will be called a set. A member of the collection is also called an element of the set. We denote by $\mathbb{R}$ the set of all numbers. To say that "x is a number" or that "x is an element of $\mathbb{R}$" amounts to the same thing. The set of n-tuples of numbers will be denoted by $\mathbb{R}^n$. Thus "X is an element of $\mathbb{R}^n$" and "X is an n-tuples" mean the same thing. Instead of saying that u is an element of a set S, we shall also frequently say that u lies in S and we write $u \in S$. If S and S' are two sets, and if every element of S' is an element of S, then we say that S' is a subset of S. Thus the set of rational numbers is a subset of the set of (real) numbers. To say that S is a subset of S' is to say that S is part of S'. To denote the fact that S is a subset of S', we write $S \subset S'$.
If S_1, S_2 are sets, then the intersection of S_1 and S_2, denoted by $S_1 \cap S_2$, is the set of elements which lie in both S_1 and S_2. The union of S_1 and S_2, denoted by $S_1 \cup S_2$, is the set of elements which lie in S_1 or S_2.

Definitions

A vector space V is a set of objects which can be added and multiplied by numbers, in such a way that the sum of two elements of V is again an element of V, the product of an element of V by a number is an element of V, and the following properties are satisfied:

VS 1 Given the elements u, v, w of V, we have

$$(u + v) + w = u + (v + w).$$

VS 2 There is an element of V denoted by O, such that

$$O + u = u + O - u.$$

for all elements u of V.

VS 3 Given an element u of V, the element $(-1)u$ is such that

$$u + (-1)u = O.$$

VS 4 For all elements u, v of V, we have

$$u + v = v + u.$$

VS 5 If c is a number, then $c(u+v)=cu+cv$.

VS 6 If a, b are two numbers, then $(a+b)v=av+bv$.

VS 7 If a, b are two numbers, then $(ab)v=a(bv)$.

VS 8 For all elements u of V, we have $1 \cdot u=u$ (1 here is the number one).

If u, v are vectors (i.e. elements of the arbitrary vector space V), then the sum

$$u+(-1)v$$

is usually written $u-v$. We also write $-v$ instead of $(-1)v$.

Fix two positive integers m, n. Let V be the set of all $m \times n$ matrices. We also denote V by $\mathrm{Mat}(m \times n)$. Then V is a vector space.
Let V be the set of all functions defined for all numbers. If f, g are two functions, then we know how to form their sum $f + g$. It is the function whose value at a number t is $f(t) + g(t)$. We also know how to multiply f by a number c. It is the function cf whose values at a number t is $cf(t)$. Then the set of functions is a vector space.

We shall use 0 to denote the number zero, and O to denote the element of any vector space V satisfying property $VS2$. We also call it zero, but there is never any possibility of confusion. We observe that this zero element O is uniquely determined by condition $VS2$. Indeed, if

$$v + w = v$$

then adding $-v$ to both sides yields

$$-v + v + w = -v + v = O,$$

and the left-hand side is just $O + w = w$, so $w = O$.
Observe that for any element v in V we have

$$\mathrm{0v{=}O.}$$

Similarly, if c is a number, then

$$\mathrm{cO{=}O.}$$

Let V be a vector space, and let W be a subset of V. Assume that W satisfies the following conditions.

(i) If v, w are elements of W, their sum $v + w$ is also an element of W.

(ii) If v is an element of W and c a number, then cv is an element of W.

(iii) The element O of V is also an element of W.

Then W itself is a vector space. Indeed, properties VS 1 through VS 8, being satisfied for all elements of V, are satisfied for the elements of W. We shall call W a subspace of V.

Let $V = \mathbb{R}^n$ and *let* W be the set of vectors in V whose last coordinate is equal to 0. Then W is a subspace of V, which we could identify with $\mathbb{R}^{n-1}$.

Let A be a vector in $\mathbb{R}^3$. Let W be the set of all elements B in $\mathbb{R}^3$ such that $B \cdot A = 0$, i.e. such that B is perpendicular to A. Then W is a subspace of $\mathbb{R}^3$.

Let $\mathrm{Sym}(n \times n)$ be the set of all symmetric $n \times n$ matrices. Then $\mathrm{Sym}(n \times n)$ is a subspace of the space of all $n \times n$ matrices.

If f, g are two continuous functions, then $f + g$ is continuous. If c is a number, then cf is continuous. The zero function is continuous. Hence, the continuous functions form a subspace of the vector space of all functions.

Let V be a vector space and let U, W be subspaces. We denote by $U \cap W$ the intersection of U and W, i.e. the set of elements which lie both in U and W. Then $U \cap W$ is a subspace.

Let U, W be subspaces of a vector space V. By $U + W$ we denote the set of all elements $u + w$ with $u \in U$ and $w \in W$. Then $U + W$ is a subspace of V, said to be generated by U and W,and called the sum of U and W.

Linear Combinations

Let V be a vector space, and let $v_1, ..., v_n$ be elements of V. We shall say that $v_1, ...v_n$ generate V if given an element $v \in V$ there exist numbers $x_1, ..., x_n$ such that

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Let $E_1, ..., E_n$ be the standard unit vectors in $\mathbb{R}^n$, so E_i has component 1 in the i-th place, and component 0 in all other places. Then $E_1, ..., E_n$ generate $\mathbb{R}^n$. Proof: given $X = (x_1, ..., x_n) \in \mathbb{R}^n$. Then

$$X = \sum_{i=1}^{n} x_i E_i,$$

so there exist numbers satisfying the condition of the definition.

Let V be an arbitrary vector space, and let $v_1, ..., v_n$ be elements of V. Let $x_1, ..., x_n$ be numbers. An expression of type

$$x_1 v_1 + \cdots + x_n v_n$$

is called a linear combination of $v_1, ..., v_n$. The numbers $x_1, ..., x_n$ are then called the coefficients of the linear combination.
The set of all linear combination of $v_1, ..., v_n$ is a subspace of V.
The subspace W consisting of all linear combination of $v_1, .., v_n$ is called the subspace generated by $v_1, ..., v_n$.

Convex Sets

Let S be a subset of a vector space V. We shall say that S is convex if given points P, Q in S then the line segment between P and Q is contained in S.
We recall that the line segment between P and Q consists of all points

$$(1-t)P + tQ \quad \text{with} \quad 0 \leq t \leq 1.$$

This gives us a simple test to determine whether a set is convex or not.

Let S be the parallelogram spanned by two vectors v_1, v_2, so S is the set of linear combinations

$$t_1 v_1 + t_2 v_2 \quad \text{with} \quad 0 \leq t_i \leq 1.$$

We wish to prove that S is convex. Let

$$P = t_1 v_1 + t_2 v_2 \quad \text{and} \quad Q = s_1 v_1 + s_2 v_2$$

be points in S. Then

$$\begin{aligned} (1-t)P + tQ &= (1-t)(t_1 v_1 + t_2 v_2) + t(s_1 v_1 + s_2 v_2) \\ &= (1-t)t_1 v_1 + (1-t)t_2 v_2 + ts_1 v_1 + ts_2 v_2 \\ &= r_1 v_1 + r_2 v_2, \end{aligned}$$

where

$$r_1 = (1-t)t_1 + ts_1 \quad \text{and} \quad r_2 = (1-t)t_2 + ts_2.$$

But we have

$$0 \leq (1-t)t_1 + ts_1 \leq (1-t) + t = 1$$

and

$$0 \leq (1-t)t_2 + ts_2 \leq (1-t) + t = 1.$$

Hence

$$(1-t)P + tQ = r_1 v_1 + r_2 v_2 \quad \text{with} \quad 0 \leq r_i \leq 1.$$

This proves that $(1-t)P + tQ$ is in the parallelogram, which is therefore convex.

Theorem

Let $P_1, ...P_n$ be points of a vector space V. Let S be the set of all linear combinations

$$t_1 P_1 + \cdots + t_n P_n$$

with $0 \leq t_i$ and $t_1 + \cdots + t_n = 1$. Then S is convex.

Proof

Let

$$P = t_1 P_1 + \cdots + t_n P_n$$

and

$$Q = s_1 P_1 + \cdots + s_n P_n$$

with $0 \leq t_i, 0 \leq s_i$, and

$$t_1 + \cdots + t_n = 1,$$
$$s_1 + \cdots + s_n = 1.$$

Proof.

Let $0 \leq t \leq 1$. Then:

$$\begin{aligned}(1-t)P + tQ &= (1-t)t_1P_1 + \cdots + (1-t)t_nP_n + ts_1P_1 + \cdots + ts_nP_n \\ &= [(1-t)t_1 + ts_1]P_1 + \cdots + [(1-t)t_n + ts_n]P_n.\end{aligned}$$

We have $0 \leq (1-t)t_i + ts_i$ for all i, and

$$\begin{aligned}&(1-t)t_1 + ts_1 + \cdots + (1-t)t_n + ts_n \\ &\quad = (1-t)(t_1 + \cdots + t_n) + t(s_1 + \cdots + s_n) \\ &\quad = (1-t) + t \\ &\quad = 1.\end{aligned}$$

This proves our theorem. □

Suppose that P_1, P_2, P_3 are three points in the plane not on a line. Then it is geometrically clear that the smallest convex set containing these three points is the triangle having these points as vertices. Thus it is natural to take as definition of a triangle the following property, valid in any vector space.

Let P_1, P_2, P_3 be three points in a vector space V, not lying on a line. Then the triangle spanned by these points is the set of all combinations

$$t_1 P_1 + t_2 P_2 + t_3 P_3 \quad \text{with} \quad 0 \leq t_i \quad \text{and} \quad t_1 + t_2 + t_3 = 1.$$

Theorem

Let $P_1, ..., P_n$ be points of a vector space V. Any convex set which contains $P_1, ..., P_n$ also contains all linear combinations

$$t_1 P_1 + \cdots + t_n P_n$$

with $0 \leq t_i$ for all i and $t_1 + \cdots + t_n = 1$.

Linear Independence

Let V be a vector space, and let $v_1, ..., v_n$ be elements of V. We shall say that $v_1, ..., v_n$ are linearly dependent if there exist numbers $a_1, ..., a_n$ not all equal to 0 such that

$$a_1 v_1 + \cdots + a_n v_n = O.$$

If there do not exists such numbers, then we say that $v_1, ..., v_n$ are linearly independent. In other words, vectors $v_1, ..., v_n$ are linearly independent if and only if the following condition is satisfied:
Let $a_1, ..., a_n$ be numbers such that

$$a_1 v_1 + \cdots + a_n v_n = O;$$

then $a_i = 0$ for all $i = 1, ..., n$.

Let $V = \mathbb{R}^n$ and consider the vectors

$$\begin{aligned} E_1 &= (1, 0, ..., 0) \\ &\vdots \\ E_n &= (0, 0, ..., 1). \end{aligned}$$

Then $E_1, ..., E_n$ are linearly independent.
The vectors $(1, 1)$ and $(-3, 2)$ are linearly independent.

If elements $v_1, \ldots, v_n$ of V generate V and in addition are linearly independent, then $\{v_1, \ldots, v_n\}$ is called a basis of V. We shall also say that the element $v_1, \ldots, v_n$ constitute or form a basis of V.
The vectors $E_1, \ldots, E_n$ form a basis of $\mathbb{R}^n$.
Any two vectors which are not parallel form a basis of $\mathbb{R}^2$. For example, the vectors $(1, 1)$ and $(-1, 2)$ form a basis of $\mathbb{R}^2$.

Theorem

Let (a, b) and (c, d) be two vectors in $\mathbb{R}^2$.

(i) *They are linearly dependent if and only if $ad - bc = 0$.*
(ii) *If they are linearly independent, then they form a basis of $\mathbb{R}^2$.*

Let V be a vector space, and let $\{v_1, ..., v_n\}$ be a basis of V. The elements of V can be represented by n-tuples relative to this basis, as follows. If an element v of V is written as a linear combination

$$v = x_1 v_1 + \cdots + x_n v_n$$

of the basis elements, then we call $(x_1, ..., x_n)$ the coordinates of v with respect to our basis, and we call x_i the i-th coordinate.

The following theorem shows that there can only be one set of coordinates for a given vector.

Theorem

Let V be a vector space. Let $v_1, ..., v_n$ be linearly independent element of V. Let $x_1, ..., x_n$ and $y_1, ..., y_n$ be numbers such that

$$x_1 v_1 + \cdots + x_n v_n = y_1 v_1 + \cdots + y_n v_n.$$

Then we must have $x_i = y_i$ for all $i = 1, ..., n$.

The theorem expresses the fact that when an element is written as a linear combination of $v_1, \ldots, v_n$, then its coefficients $x_1, \ldots, x_n$ are uniquely determined. This is true only when $v_1, \ldots, v_n$ are linearly independent.
Find the coordinates of $(1, 0)$ with respect to the two vectors $(1, 1)$ and $(-1, 2)$.
The two functions e^t, e^{2t} are linearly independent.
Let V be the vector space of all functions of a variable t. Let $f_1, \ldots, f_n$ be n functions. To say that they are linearly dependent is to say that there exist n numbers $a_1, \ldots, a_n$ not all equal to 0 such that

$$a_1 f_1(t) + \cdots + a_n f_n(t) = 0$$

for all values of t.

When dealing with two vectors v, w there is another convenient way of expressing linear independence.

Theorem

Let v, w be elements of a vector space V. They are linearly dependent if and only if one of them is a scalar multiple of the other, i.e. there is a number $c \neq 0$ such that we have $v = cw$ or $w = cv$.

Dimension

We ask the question: Can we find three linearly independent elements in $\mathbb{R}^2$? For instance, are the elements

$$A = (1, 2), \quad B = (-5, 7), \quad C = (10, 4)$$

linearly independent? If you write down the linear equations expressing the relation

$$xA + yB + zC = O,$$

you will find that you can solve them for x, y, z not equal to 0. Namely, these equations are

$$\begin{aligned} x - 5y + 10z &= 0, \\ 2x + 7y + 4z &= 0. \end{aligned}$$

This is a system of two homogeneous equations in three unknowns, and we can find a non-trivial solution (x, y, z) not all equal to zero. Hence, A, B, C are linearly dependent.

We shall see in a moment that this is a general phenomenon. In $\mathbb{R}^n$, we cannot find more than n linearly independent vectors. Furthermore, we shall see that any n linearly independent elements of $\mathbb{R}^n$ must generate $\mathbb{R}^n$, and hence form a basis. Finally, we shall also see that if one basis of a vector space has n elements, and another basis has m elements, then $m = n$. In short, two bases must have the same number of elements. This property will allow us to define the dimension of a vector space as the number of elements in any basis. We now develop these ideas systematically.

Theorem

Let V be a vector space, and let $\{v_1, ..., v_m\}$ generate V. Let $w_1, ..., w_n$ be elements of V and assume that $n > m$. Then $w_1, ..., w_n$ are linearly dependent.

Proof

Sine $\{v_1, ..., v_m\}$ generate V, there exist numbers (a_{ij}) such that we can write

$$\begin{array}{ccccc} w_1 & = & a_{11}v_1 & + \cdots + & a_{m1}v_m \\ \vdots & & \vdots & & \vdots \\ w_n & = & a_{1n}v_1 & + \cdots + & a_{mn}v_m. \end{array}$$

If $x_1, ..., x_n$ are numbers, then

$$\begin{aligned} & x_1 w_1 + \cdots + x_n w_n \\ & \quad = (x_1 a_{11} + \cdots + x_n a_{1n}) v_1 + \cdots + (x_1 a_{m1} + \cdots + x_n a_{mn}) v_m. \end{aligned}$$

Proof.

Then the system of equations

$$\begin{array}{rcl} x_1 a_{11} & + \cdots + x_n a_{1n} & = & 0 \\ \vdots & & \vdots & \\ x_1 a_{m1} & + \cdots + x_n a_{mn} & = & 0 \end{array}$$

has a non-trivial solution, because $n > m$. In view of the preceding remark, such a solution $(x_1, ..., x_n)$ is such that

$$x_1 w_1 + \cdots + x_n w_n = 0$$

as desired. □

Theorem

Let V be a vector space and suppose that one basis has n elements, and another basis has m elements. Then $m = n$.

Proof.

We apply the previous theorem to the two bases. It implies that both alternatives $n > m$ and $m > n$ are impossible, and hence $m = n$. □

Let V be a vector space having a basis consisting of n elements. We shall say that n is the dimension of V. If V consists of O alone, then V does not have a basis, and we shall say that V has dimension 0.

We may now reformulate the definitions of a line and a plane in an arbitrary vector space V. A line passing through the origin is simply a one-dimensional subspace. A plane passing through the origin is simply a two-dimensional subspace.

We may now reformulate the definitions of a line and a plane in an arbitrary vector space V. A line passing through the origin is simply a one-dimensional subspace. A plane passing through the origin is simply a two-dimensional subspace.

An arbitrary line is obtained as the translation of a one-dimensional subspace. An arbitrary plane is obtained as the translation of a two-dimensional subspace. When a basis $\{v_1\}$ has been selected for a one-dimensional space, then the points on a line are expressed in the usual form

$$P + t_1 v_1 \text{ with all possible numbers } t_1.$$

When a basis $\{v_1, v_2\}$ has been selected for a two-dimensional space, then the points on a plane are expressed in the form

$$P + t_1 v_1 + t_2 v_2 \text{ with all possible numbers } t_1, t_2.$$

Let $\{v_1, ..., v_n\}$ be a set of elements of a vector space V. Let r be a positive integer $\leq n$. We shall say that $\{v_1, ..., v_r\}$ is a maximal subset of linearly independent elements if $v_1, ..., v_r$ are linearly independent, and if in addition, given any v_i with $i > r$, the elements $v_1, ..., v_r, v_i$ are linearly dependent.

The next theorem gives us a useful criterion to determine when a set of elements of a vector space is a basis.

Theorem

Let $\{v_1, ..., v_n\}$ be a set of generators of a vector space V. Let $\{v_1, ..., v_r\}$ be a maximal subset of linearly independent elements. Then $\{v_1, ..., v_r\}$ is a basis of V.

Proof.

We must prove that $v_1, \ldots, v_r$ generate V. We shall first prove that each v_i (for $i > r$) is a linear combination of $v_1, \ldots, v_r$. By hypothesis, given v_i, there exists numbers $x_1, \ldots, x_r, y$ not all 0 such that

$$x_1 v_1 + \cdots + x_r v_r + y v_i = O.$$

Furthermore, $y \neq 0$, because otherwise, we would have a relation of linear dependence for $v_1, \ldots, v_r$. Hence we can solve for v_i, namely

$$v_i = \frac{x_1}{-y} v_1 + \cdots + \frac{x_r}{-y} v_r,$$

thereby showing that v_i is a linear combination of $v_1, \ldots, v_r$.
Next, let v be any element of V. There exist numbers $c_1, \ldots, c_n$ such that

$$v = c_1 v_1 + \cdots + c_n v_n.$$

In this relation, we can replace each $v_i (i > r)$ by a linear combination of $v_1, \ldots, v_r$. If we do this, and then collect terms, we find that we have expressed v as a linear combination of $v_1, \ldots, v_r$. This proves that $v_1, \ldots, v_r$ generate V, and hence form a basis of V. □

We shall now give criteria which allow us to tell when elements of a vector space constitute a basis.

Let $v_1, \ldots, v_n$ be linearly independent elements of a vector space V. We shall say that they form a maximal set of linearly independent elements of V if given any element w of V, the elements $w, v_1, \ldots, v_n$ are linearly dependent.

Theorem

Let V be a vector space, and $\{v_1, \ldots, v_n\}$ a maximal set of linearly independent elements of V. Then $\{v_1, \ldots, v_n\}$ is a basis of V.

Theorem

Let V be a vector space of dimension n, and let $v_1, ..., v_n$ be linearly independent elements of V. Then $v_1, ..., v_n$ constitute a basis of V.

Theorem

Let V be a vector space of dimension n and let W be a subspace, also of dimension n. Then $W = V$.

Theorem

Let V be a vector space of dimension n. Let r be a positive integer with $r < n$, and let $v_1, ..., v_r$ be linearly independent elements of V. Then one can find elements $v_{r+1}, ..., v_n$ such that

$$\{v_1, ..., v_n\}$$

is a basis of V.

Theorem

Let V be a vector space having a basis consisting of n elements. Let W be a subspace which does not consist of O alone. Then W has a basis, and the dimension of W is $\leq n$.

Proof.

Let w_1 be a non-zero element of W. If $\{w_1\}$ is not a maximal set of linearly independent elements of W, we can find an element w_2 of W such that w_1, w_2 are linearly independent. Proceeding in this manner, one element at a time, there must be an integer $m \leq n$ such that we can find linearly independent elements $w_1, w_2, ..., w_m$, and such that

$$\{w_1, ..., w_m\}$$

is a maximal set of linearly independent elements of W. We conclude that $\{w_1, ..., w_m\}$ is a basis for W. □

The Rank of a Matrix

Let

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

be an $m \times n$ matrix. The columns of A generate a vector space, which is a subspace of $\mathbb{R}^m$. The dimension of that subspace is called the column rank of A. The column rank is equal to the maximum number of linearly independent columns. Similarly, the rows of A generate a subspace of $\mathbb{R}^n$, and the dimension of this subspace is called the row rank. The row rank is equal to the maximum number of linearly independent rows. We shall prove below that these two ranks are equal to each other.

We define the row space of A to be the subspace generated by the rows of A. We define the column space of A to be the subspace generated by the columns.

Consider the following operations on the rows of a matrix.

Row 1. Adding a scalar multiple of one row to another.

Row 2. Interchanging rows.

Row 3. Multiplying one row by a non-zero scalar.

These are called the row operations (sometimes, the elementary row operations). We have similar operations for columns, which will be denoted by Co 1, Co2 Co3 respectively. We shall study the effect of these operations on the ranks.

First observe that each one of the above operations has an inverse operation in the sense that by performing similar operations we can revert to the original matrix. For instance, let us change a matrix A by adding c times the second row to the first. We obtain a new matrix B whose rows are

$$B_1 = A_1 + cA_2, A_2, ..., A_m.$$

If we now add $-cA_2$ to the first row of B, we get back A_1. A similar argument can be applied to any two rows.
If we interchange two rows, then interchange them again, we revert to the original matrix.
If we multiply a row by a number $c \neq 0$, then multiplying again by c^{-1} yields the original row.

Theorem

Row and column operations do not change the row rank of a matrix, nor do they change the column rank.

Theorem

Let A be a matrix of row rank r. By a succession of row and column operations, the matrix can be transformed to the matrix having components equal to 1 on the diagonal of the first r rows and columns, and 0 everywhere else.

Since we have proved that the row rank is equal to the column rank, we can now omit "row" or "column" and just speak of the rank of a matrix. Thus by definition the rank of a matrix is equal to the dimension of the space generated by the rows.

Find the rank of the matrix

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

Find the rank of the matrix

$$\begin{pmatrix} 1 & 2 & -3 \\ 2 & 1 & 0 \\ -2 & -1 & 3 \\ -1 & 4 & -2 \end{pmatrix}.$$