

On the Bases of a Vector Space

Kim Young-sinn

Contents

I	Introduction	§ 3. Orthogonality
II	Main Theorems	§ 4. Continuity of a linear map
	§ 1. Vector space	III Conclusions
	§ 2. Bases	Summary

I. Introduction

In modern mathematics, especially in algebra, the structure of the (algebraic) number system is very important.

I am going to study the following structures :

- (1) field,
- (2) vector space over a field,
- (3) subspaces of a vector space,
- (4) basis of a vector space
(especially, orthogonal basis),
- (5) The properties of orthogonality and
- (6) The convergence in a vector space and continuity of a linear map.

II. Main Theorems

§ 1. Vector space

A set K of numbers with the two operation (addition and multiplication) is called a field if it has the following properties:

(A) To every pair, α and β , of K there corresponds a number $\alpha + \beta$ in K , called the sum of α and β , in such a way that

- (1) addition is commutative, $\alpha + \beta = \beta + \alpha$,
- (2) addition is associative, $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$,
- (3) there exists a unique number o (called zero) such that $\alpha + o = \alpha$ for every number α of K , and
- (4) to every number α there corresponds a unique number $-\alpha$ such that $\alpha + (-\alpha) = o$

(B) To every pair, α and β , of K there corresponds a number $\alpha \cdot \beta$ in K , called the multiplication of α and β , in such a way that

- (1) multiplication is commutative, $\alpha\beta = \beta\alpha$,
- (2) multiplication is associative, $\alpha(\beta\gamma) = (\alpha\beta)\gamma$,
- (3) there exists a unique non-zero number 1 (called one) such that $\alpha \cdot 1 = \alpha$ for every α of K , and
- (4) to every non-zero number α of K , there corresponds a unique number α^{-1} (or $\frac{1}{\alpha}$) such that $\alpha \cdot \alpha^{-1} = 1$

(C) Multiplication is distributive with respect to addition, $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

Sometimes, the elements of K are called the scalars.

Thus, for example, the set of all rational numbers Q (with the ordinary definitions of addition and multiplication) is a field, and the same is true of the set R of all real numbers and the set C of all complex numbers.

An example of finite field:

Let Z be the set of all integers and p be a prime number, then the residue classes $Z/(p)$ (with ordinary addition and multiplication) is a field, which has only p elements.

If S and S' are sets, and if every element of S' is an element of S , then we say that " S' is a subset of S ," and denote by $S' \subset S$; furthermore, $S' \neq S$, then S' is called a proper subset of S .

Let K, L be fields, and suppose that K is contained in L (i. e. that K is a subset of L). Then we shall say that K is a subfield of L . Thus R is a subfield of C , and Q is a subfield of R and C .

We now come to the basic concepts of this note. (We assume that a particular field K is given.)

DEFINITION. A vector space is a set V of elements called vectors satisfying the following axioms.

(A) To every pair, u and v , of vectors in V there corresponds a unique vector $u+v$, called the addition of u and v , in such a way that

- (1) addition is commutative, $u+v=v+u$,
- (2) addition is associative, $u+(v+w)=(u+v)+w$,
- (3) there exists in V a unique vector O (called the zero vector) such that $u+O=u$ for every vector u , and
- (4) to every vector u in V there is a unique vector $(-u)$ in V such that $u+(-u)=O$.

(B) To every pair, α and u (α is in K and u is in V) there corresponds a vector αu in V (called the scalar multiplication) in such a way that

- (1) multiplication by scalars is associative, $\alpha(\beta u)=(\alpha\beta)u$, and
- (2) $1 \cdot u = u$ for every vector u .

(C) (1) multiplication by scalars is distributive with respect to vector addition,

$$\alpha(u+v) = \alpha u + \alpha v, \text{ and}$$

(2) multiplication by vectors is distributive with respect to scalar addition,

$$(\alpha + \beta)u = \alpha u + \beta u$$

The relation between a vector space V and the underlying field K is usually described by saying " V is a vector space over K ".

Thus, if K is the field R of real numbers, V is called a real vector space; similarly if K is Q or K is C , we speak of rational vector spaces or complex vector spaces.

Examples of vector spaces

(1) Let $C'(=C)$ be the set of all complex numbers: if we interpret $u+v$ and au as ordinary complex numerical addition and multiplication, C becomes a complex vector space.

(2) Let $C^n(n=1, 2, \dots, n)$ be the set of all n -tuples of complex numbers.

If $u = (\alpha_1, \dots, \alpha_n)$ and $v = (\beta_1, \dots, \beta_n)$ are elements of C^n , we write (by definition)

$$U+V = (\alpha_1+\beta_1, \dots, \alpha_n+\beta_n)$$

$$\gamma U = (\gamma\alpha_1, \dots, \gamma\alpha_n)$$

$$O = (0, \dots, 0)$$

$$(-U) = (-\alpha_1, \dots, -\alpha_n)$$

It is easy to verify that C^n is a complex vector space; it is some times called n -dimensional complex coordinate space.

(3) A close relative of C^n is R^n of all n -tuples of real numbers. With the same formal definitions of addition and scalar multiplication as for C^n , except that we consider only real scalars α , the space R^n is a real vector space; it will be called the n -dimensional coordinate space.

(4) Let F be the set of all continuous functions on the interval $[0, 1]$, then F is a vector space over R .

DEFINITION. Let W be a subset of a vector space V , which satisfies the following properties:

- (1) If u, v are elements of W , their sum $u+v$ is also in W .
- (2) If v is an element of W and α is an element of K , then αv is also in W .
- (3) The element 0 of V is also in W .

Then W is called a subspace of V , which is also a vector space.

The above definition is equivalent to the

DEFINITION. A non-empty subset W of a vector space V is a subspace (or a linear manifold) if for every u, v of W , their linear combination

$$\alpha u + \beta v$$

is contained in W .

Two special examples of subspaces are:

- (1) The set $\{0\}$ consisting of the zero vector only,
- (2) The whole space V ,

and the other subspaces are proper ones.

THEOREM 1. The intersection of any collection of subspaces is a subspace.

PROOF. If we use an index i to tell apart the members of the collection, so that the given subspaces of the vector space V are W_i , let us write

$$W = \bigcap_i W_i$$

Since every W_i contains 0 , so does W , and therefore W is not empty. If u and v belong to W (that is, to all W_i), their linear combination $\alpha u + \beta v$ belongs to all W_i , and hence to W .

Therefore W is a subspace of V .

Q. E. D.

§ 2. Linear Basis

DEFINITION. A finite set $\{v_i\}$ of vectors is linearly dependent if there exists a corresponding set $\{x_i\}$ of scalars, not all zero, such that

$$\sum_i x_i v_i = 0$$

If, on the other hand, $\sum x_i v_i = 0$ implies that $x_i = 0$ for each i , the set $\{v_i\}$ is linearly independent.

In R^3 , $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ are linearly independent. However, $a = (1, 0, 0)$, $b = (0, 1, 0)$, $c = (1, 2, 0)$ are linearly dependent.

For convention let us describe "iff" instead of "if and only if".

THEOREM 2. The set of non-zero vectors $\{v_1, \dots, v_n\}$ are linearly dependent iff some v_k , $2 \leq k \leq n$, is a linear combination of the preceding ones.

PROOF. Let us suppose that the vectors $\{v_1, \dots, v_n\}$ are linearly dependent, and let k be the first integer between 2 and n for which $\{v_1, \dots, v_k\}$ are linearly dependent. Then

$$x_1 v_1 + \dots + x_k v_k = 0$$

for a suitable set of $\{x_i\}$ (not all zero); moreover, whatever the x 's, we can not have $x_k = 0$, for then we should have a linear dependent relation among $\{v_1, \dots, v_{k-1}\}$, contrary to the definition of k . Hence

$$x_k v_k = -x_1 v_1 - \dots - x_{k-1} v_{k-1}.$$

Dividing both sides by x_k , we have

$$v_k = -\frac{x_1}{x_k} v_1 - \dots - \frac{x_{k-1}}{x_k} v_{k-1}$$

, which is a linear combination of preceding ones.

Q. E. D.

DEFINITION. A (linear) basis (or a coordinate system) in a vector space V is a set \mathbf{B} of linearly independent vectors such that every vector in V is a linear combination of elements of \mathbf{B} . A vector space V is finite dimensional if it has a finite basis.

If $e_1=(1, 0, 0)$, $e_2=(0, 1, 0)$, $e_3=(0, 0, 1)$, then $B=\{e_1, e_2, e_3\}$ is a basis of R^3 , which is a 3-dimensional vector space.

However, let F be the set of all continuous functions on $[-\pi, \pi]$, then F is a vector space. If $\mathbf{B}=\{\dots, \sin 2t, \sin t, 1, \cos t, \cos 2t, \dots\}$, then \mathbf{B} is a basis of F , which is an infinite-dimensional.

THEOREM 3. In a finite-dimensional vector space V , with basis $\{v_1, \dots, v_n\}$, every v of V is written in the form

$$v = \sum x_i v_i$$

, then the x 's are uniquely determined by v .

PROOF. If $v = \sum y_i v_i$, then by subtracting we have

$$\sum (\alpha_i - \beta_i) v_i = 0$$

Since the v 's are linearly independent, this implies that $\alpha_i - \beta_i = 0$ for $i=1, \dots, n$; in other words, the α 's are the same as the corresponding β 's.

Q. E. D.

THEOREM 4. If V is a finite-dimensional vector space and if $\{v_1, \dots, v_m\}$ is any set of linearly independent vectors in V , then, unless the v 's already form a basis, we can find vectors $\{v_{m+1}, \dots, v_n\}$ so that the totality of the v 's, i. e. $\{v_1, \dots, v_m, \dots, v_n\}$ is a basis of V . In other words, every linearly independent set can be extended to a basis.

PROOF. Since V is finite-dimensional, it has a finite basis, say $\{u_1, \dots, u_n\}$.

We consider the set S of vectors:

$$S = \{v_1, \dots, v_m, u_1, \dots, u_n\}$$

Let us apply Theorem 2 to this set several times in succession.

In the first place, (since the v 's are linear combinations of the u 's) the set S is linearly independent. Hence, some vector of S is a linear combination of the preceding ones.

Let w_1 be the first such vector.

Since the v 's are linearly independent, w_1 is different from any

$$v_i (i=1, \dots, m),$$

so that w_1 is equal to some u_k , say $w_1 = u_k$.

We consider the new set S_1 of vectors:

$$S_1 = \{v_1, \dots, v_m; u_1, \dots, u_{k-1}, u_{k+1}, \dots, u_n\}$$

Since $\{u_1, \dots, u_k, \dots, u_n\}$ is a basis and we may express u_k by means of $\{v_1, \dots, v_m, u_1, \dots, u_{k-1}\}$, we observe that every vector in V is a linear combination of vectors in S_1 . If S_1 is linearly independent, we are done. If S_1 is not linearly independent, we can repeat the above procedure until the remaining set to be linearly independent. The remaining set is the required basis which contains obviously the linearly independent set $\{v_1, \dots, v_m\}$.

Q. E. D.

Let V be a vector space over a field K , and let $\mathbf{B} = \{v_1, \dots, v_n\}$ be a basis. If v is any element of V , then the set

$$\{v_1, \dots, v_n; v\}$$

is linearly dependent. Hence, the number of a basis, which are linearly independent in V and generates V , is very important.

DEFINITION. The dimension of a vector space V is the number of elements in a basis for V .

Remark. If the number of a basis of V is infinite, then the space V is called infinite dimensional, and if finite, finite dimensional.

THEOREM 5. The number of elements in any basis of a finite dimensional space V is the same as in any other basis.

PROOF. The proof of this Theorem is a slight refinement of the method used in Theorem 4.

Let $\mathbf{B}=\{u_1, \dots, u_n\}$ and $\mathbf{B}'=\{v_1, \dots, v_m\}$ be two finite sets of vectors, and every vector in V is a linear combination of \mathbf{B} (but not necessarily linearly independent), and \mathbf{B}' are linearly independent (but not necessarily generates V)

That is, \mathbf{B} generates V and \mathbf{B}' are independent in V . Let us consider the set S of vectors:

$$S=\{v_m, u_1, u_2, \dots, u_n\}$$

Since every vector of V is a linear combination of \mathbf{B} , S is linearly dependent.

Reasoning just as Theorem 4, we obtain a set S_1 of vectors:

$$S_1=\{v_m, u_1, \dots, u_{k-1}, u_{k+1}, \dots, u_n\}$$

S_1 has the same property to \mathbf{B} .

Now we write v_{m-1} in front of the vectors of S_1 and apply the same argument.

Repeating m times in this way, we have the set

$$S_0=\{v_1, v_2, \dots, v_m; u_1', \dots, u_i'\},$$

where every u_i' is some u_k in \mathbf{B} .

Hence, $n \geq m$.

Consequently, if both \mathbf{B} and \mathbf{B}' are bases (with the above two properties), then $n \geq m$ and $m \geq n$. Hence, $m = n$.

Q. E. D.

COROLLARY 1. An n -dimensional vector space V has a basis, which has just n independent elements. And a set of n independent vectors in V forms a basis.

PROOF. The result is trivial from the above Theorem 4 and 5.

Q. E. D.

COROLLARY 2. Every set of $n+1$ vectors in an n -dimensional vector space V is linearly dependent.

PROOF. Since a basis \mathbf{B} of the n -dimensional vector space V has just n elements, every vector of V is expressed as a linear combination of \mathbf{B} . Hence, $n+1$ vectors are linearly dependent.

Q. E. D.

As an application of the notion of linear basis, or coordinate system, we shall see that every n -dimensional vector space V over K is essentially the same as (is isomorphic to) K^n .

DEFINITION. Let V and V' be two vector spaces over K . A linear mapping (map, transformation)

$$F : V \longrightarrow V'$$

is a mapping such that

$$F(\alpha u + \beta v) = \alpha F(u) + \beta F(v) \quad (\text{where } u, v \text{ are in } V)$$

If $V' = V$, we call the mapping F "an operator on V ."

DEFINITION. Two vector spaces V and V' (over the same field K) are isomorphic if there is a one-to-one linear transformation between them.

THEOREM 6. Every n -dimensional vector space V over a field K is isomorphic to K^n .

PROOF. Let $\{v_1, \dots, v_n\}$ be any basis in V . Each v in V can be written in the form

$$v = x_1v_1 + x_2v_2 + \dots + x_nv_n,$$

and we know that the scalars x_1, \dots, x_n are uniquely determined by v . We consider the one-to-one correspondence

$$v \sim (x_1, x_2, \dots, x_n)$$

between V and K^n . If $u = y_1v_1 + \dots + y_nv_n$, then

$$\alpha v + \beta u = (\alpha x_1 + \beta y_1)v_1 + \dots + (\alpha x_n + \beta y_n)v_n$$

The above correspondence establishes the desired isomorphism.

Q. E. D.

Suppose that S is an arbitrary set of vectors (not necessarily a subspace) in a vector space V . There certainly exist subspaces W_i containing every element of S (i. e., such that $S \subset W_i$). For example, the whole space V is such a subspace.

Let W be the intersection of all the subspaces containing S . By Theorem 1, W is the smallest subspace containing S .

DEFINITION. Let S be a subset of a vector space V . The smallest subspace W containing S is called the subspace generated (spanned) by S , and W is called the span of S .

THEOREM 7. If S is any set of vectors in a vector space V and if W is the span of S , then W is the same as the set of all linear combinations of elements of S .

PROOF. It is clear that a linear combination of linear combinations of elements of S may again be written as a linear combination of elements of S . Hence the set of all linear combinations of elements of S is a subspace containing S , and moreover, this subspace must also contain W . On the other hand, W contains S and is a subspace.

Hence, W contains all linear combinations of elements of S . (i. e., If $W = \text{the span of } S$, then the span of $W = W$).

Q. E. D.

THEOREM 8. If H and K are any two subspaces and W is the subspace spanned by H and K together, then W is the same as the set of all vectors of the form $u+v$ with u of H and v of K .

PROOF. For u_1, u_2 of H and v_1, v_2 of K , let $u_1+v_1=w_1$ and $u_2+v_2=w_2$, then

$$\alpha w_1 + \beta w_2 = \alpha(u_1+v_1) + \beta(u_2+v_2) = (\alpha u_1 + \beta u_2) + (\alpha v_1 + \beta v_2)$$

is in W (because, $\alpha u_1 + \beta u_2$ is in H and $\alpha v_1 + \beta v_2$ is in K).

Hence, W is a subspace of V .

Q. E. D.

We shall use the notation $H+K$ for the subspace W spanned by H and K .

DEFINITION. A subspace K of a vector space is a complement of a subspace H if

$$H \cap K = \{0\} \quad \text{and} \quad H + K = V.$$

THEOREM 9. A subspace W in an n -dimensional vector space V is a vector space of dimension $\leq n$.

PROOF. Since every set of $n+1$ vectors in V is linearly dependent, the same is true in W .

Hence, the number of elements in each basis of W is $\leq n$.

Q. E. D.

From the above Theorem, We have a very essential

COROLLARY. Given any m -dimensional subspace W in an n -dimensional vector space V , we can find a basis $\{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$ in V such that $\{v_1, \dots, v_m\}$ is a basis of W .

PROOF. Since the dimension of W (WCV) is m , there is a basis $\{v_1, \dots, v_m\}$ of W .

Since $\{v_1, \dots, v_m\}$ are linearly independent in W , and so in V . By the above theorem $n \geq m$.

If $m=n$, then there is nothing to talk about.

If $m < n$, then there is a vector v_{m+1} such that

$$v_{m+1} \text{ is not in } W \text{ and } v_{m+1} \text{ is in } V.$$

Let us consider a set

$$S_1 = \{v_1, \dots, v_m, v_{m+1}\}$$

and the span of $S_1(W_1)$.

However, $\dim V \geq \dim W_1$ (i. e., $n \geq m+1$).

Repeating the same process, we can finally get

$$S_0 = \{v_1, \dots, v_m, v_{m+1}, \dots, v_n\},$$

which is independent in V , and hence a basis of V .

Q. E. D.

§ 3. Orthogonality

Let us first inspect $V = \mathbb{R}^2$. If $u = (x_1, x_2)$ and $v = (y_1, y_2)$ are any two points (*vectors*)

in R^2 , the usual formula for the distance between u and v , or the length of the segment joining u and v , is

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

It is convenient to introduce the notation

$$\|u\| = \sqrt{x_1^2 + x_2^2}$$

for the distance from u to the origin $O=(0,0)$.

In this notation the distance between u and v becomes

$$\|u - v\|$$

Furthermore, the cosine of the angle between the vectors u and v is

$$\frac{x_1 y_1 + x_2 y_2}{\sqrt{x_1^2 + x_2^2} \sqrt{y_1^2 + y_2^2}} = \frac{(u, v)}{\|u\| \|v\|}$$

The important properties of (u, v) , (considered as a numerical function of the pair of vectors u and v) are the following ones:

DEFINITION. An inner product (or scalar product) in a (real or complex) vector space is a (real or complex) valued function of the ordered pair of vectors u and v , such that

- (1) $(u, v) = \overline{(v, u)}$
- (2) $(\alpha_1 u_1 + \alpha_2 u_2, v) = \alpha_1 (u_1, v) + \alpha_2 (u_2, v)$
- (3) $(u, u) \geq 0$ and $(u, u) = 0$ iff $u = 0$.

An inner product space is a vector space with an inner product (scalar product).

We observe that in the case of real vector space, the conjugation in (1) may be ignored. In any case, however, real or complex, (1) implies that (u, u) is always positive real, so that the inequality in (3) makes sense. In an inner product space we shall use the notation the length of a vector.

DEFINITION. The number $\|u\| = \sqrt{\langle u, u \rangle}$ is called the norm (or the length) of the vector u .

A real inner product space is sometimes called a Euclidean space, and its complex analogue is called a unitary space.

As an example of unitary space we consider C^n : for $u=(x_1, \dots, x_n)$ and $v=(y_1, \dots, y_n)$,

$$\langle u, v \rangle = \sum_i x_i \overline{y_i}$$

and, let F be the set of all continuous complex valued functions on $[0, 1]$, f and g are in F , then their inner product is

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt.$$

LEMMA 1. In a unitary space we have

$$(1) \langle u, \alpha_1 u_1 + \alpha_2 u_2 \rangle = \overline{\alpha_1} \langle u, u_1 \rangle + \overline{\alpha_2} \langle u, u_2 \rangle$$

$$(2) \|\alpha u\| = |\alpha| \|u\|.$$

PROOF. (1) $\langle u, \alpha_1 v_1 + \alpha_2 v_2 \rangle = \overline{(\alpha_1 v_1 + \alpha_2 v_2, u)}$

$$= \overline{(\alpha_1 v_1, u) + (\alpha_2 v_2, u)}$$

$$= \overline{\alpha_1 (v_1, u) + \alpha_2 (v_2, u)}$$

$$= \overline{\alpha_1} \overline{(v_1, u)} + \overline{\alpha_2} \overline{(v_2, u)}$$

$$= \overline{\alpha_1} \langle u, v_1 \rangle + \overline{\alpha_2} \langle u, v_2 \rangle$$

$$(2) \|\alpha u\|^2 = \langle \alpha u, \alpha u \rangle = \alpha \overline{\alpha} \langle u, u \rangle = |\alpha|^2 \|u\|^2$$

Hence, $\|\alpha u\| = |\alpha| \|u\|.$

Q. E. D.

I suppose that the most important relation among the vectors of an inner product space is orthogonality.

DEFINITION. The vectors u and v are called orthogonal (or perpendicular) if $\langle u, v \rangle = 0$. Two subspaces are called orthogonal if every vector in each subspace is orthogonal to every vector in the other subspace.

A set S of vectors is orthogonal if whenever both u and v are in S it follows that $(u, v) = 0$ or $(u, v) \neq 0$ according as $u \neq kv$ or $u = kv$.

If for every u of S $\|u\| = 1$, then the set is called orthonormal set. (In this case, let S be finite, say $S = \{u_1, \dots, u_n\}$, we have $(u_i, u_j) = \delta_{ij}$.)

We call an orthonormal set complete if it is not contained in any larger orthonormal set.

LEMMA 2. An orthonormal set is linearly independent.

PROOF. If $\{v_1, \dots, v_k\}$ is any finite subset of an orthonormal set S , then

$$\sum_{i=1}^k \alpha_i v_i = 0.$$

implies that

$$0 = (\sum \alpha_i v_i, v_j) = \sum \alpha_i (v_i, v_j) = \sum \alpha_i \delta_{ij} = \alpha_j$$

Hence, a linear combination of v 's can vanish only if all of the coefficients vanish.

Q. E. D.

We need some notations. If W is any set of vectors in an inner product space V , we denote by W^\perp the set of all vectors in V that are orthogonal to every vector in W .

LEMMA 3. W^\perp is a subspace of V (whether W is a subspace or not), and W is in $W^{\perp\perp}$ which is the span of W .

PROOF. If W is the whole space V , there is nothing to talk about. If W is a proper subset then for every u of W , there exist v_1, v_2 in W^\perp ($W^\perp \subset V$) such that

$$(u, v_1) = 0 \text{ and } (u, v_2) = 0$$

$$(u, \alpha v_1 + \beta v_2) = \bar{\alpha}(u, v_1) + \bar{\beta}(u, v_2) = 0;$$

hence the fact that v_1, v_2 are in W^\perp implies $\alpha v_1 + \beta v_2$ is in W^\perp .

Therefore, W^\perp is a subspace of V . Furthermore, $W^{\perp\perp}$ is a subspace of V which is perpendicular to W^\perp . Let u be in W , then u is in $W^{\perp\perp}$ and hence

$$u \text{ is in } (W^\perp)^\perp \text{ i. e., } u \text{ is in } W^{\perp\perp}$$

Hence, $W \subset W^{\perp\perp}$.

Q. E. D.

THEOREM 10. (Bessel's inequality)

If $S = \{v_1, \dots, v_n\}$ is any finite orthonormal set in an inner product space, if v is any vector and if $\alpha_i = (v, v_i)$, then

$$\sum_i |\alpha_i|^2 \leq \|v\|^2$$

The vector

$$v' = v - \sum \alpha_i v_i$$

is orthogonal to every v_i and, consequently, to the subspace spanned by S .

PROOF. For the first assertion:

$$\begin{aligned} 0 \leq \|v'\|^2 &= (v', v') \\ &= (v - \sum \alpha_i v_i, v - \sum \alpha_i v_i) \\ &= (v, v) - (v, \sum \alpha_i v_i) - (\sum \alpha_i v_i, v) + \sum \alpha_i \alpha_j (v_i, v_j) \\ &= \|v\|^2 - \sum \bar{\alpha}_i (v, v_i) - \sum \alpha_i (v_i, v) + \sum \alpha_i \bar{\alpha}_j \delta_{ij} = \|v\|^2 - \sum \bar{\alpha}_i \alpha_i - \sum \alpha_i \bar{\alpha}_i + \sum \alpha_i \bar{\alpha}_i \\ &= \|v\|^2 - \sum \alpha_i \bar{\alpha}_i \\ &= \|v\|^2 - \sum |\alpha_i|^2 \end{aligned}$$

Hence, $\|v\|^2 \geq \sum |\alpha_i|^2$

For the second assertion:

$$\begin{aligned} (v', v_i) &= (v - \sum \alpha_j v_j, v_i) = (v, v_i) - \sum \alpha_j (v_j, v_i) \\ &= \alpha_i - \sum \alpha_j \delta_{ij} = \alpha_i - \alpha_i = 0 \end{aligned}$$

Hence, v' is orthogonal to the span of S .

Q. E. D.

THEOREM 11. If S is any finite orthonormal set in an inner product space V , the following six conditions on S are equivalent to each other. ($\dim V = n$)

- (1) The orthonormal set S is complete.
- (2) If $(v, v_i) = 0$ for $i=1, \dots, n$, then $v=0$.
- (3) The subspace spanned by S is the whole space V .
- (4) If v is in V , then $v = \sum_i (v, v_i)v_i$.
- (5) If u and v are in V , then

$$(u, v) = \sum_i (u, v_i)(v_i, v).$$

- (6) If v is in V , then

$$\|v\|^2 = \sum_i |(v, v_i)|^2.$$

PROOF. Let us try to show $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5) \Rightarrow (6) \Rightarrow (1)$.

Thus we first assume (1).

(1) \Rightarrow (2). Let us show the contrapositive case, *i.e.*, $\text{not}(2) \Rightarrow \text{not}(1)$.

If $(v, v_i) = 0$ for all i and $v \neq 0$, then we may adjoin

$$\frac{v}{\|v\|}$$

to S , and thus

$$S' = \{v_1, \dots, v_n; \frac{v}{\|v\|}\}$$

is an orthonormal set larger than S .

Hence, S is not complete.

Since the contrapositive proposition of true one is true, $(1) \Rightarrow (2)$

(2) \Rightarrow (3). Let us show that the contrapositive proposition of $(2) \Rightarrow (3)$ is true. *i.e.*, $\text{not}(3) \Rightarrow \text{not}(2)$. If there is a v in V that is not a linear combination of the v_i , then by Theorem 10,

$$v' = v - \sum_i (v, v_i)v_i$$

is different from 0 and orthogonal to each v_i . Hence

$$v \neq \sum_i (v, v_i)v_i$$

Hence, (2) \Rightarrow (3) is right.

(3) \Rightarrow (4). If every v has the form $v = \sum \alpha_i v_i$, then

$$(v, v_i) = (\sum \alpha_j v_j, v_i) = \sum \alpha_j (v_j, v_i) = \sum \alpha_j \delta_{ji} = \alpha_i.$$

Hence, $v = \sum_i (v, v_i)v_i$.

(4) \Rightarrow (5). If $u = \sum \alpha_i v_i$ and $v = \sum \beta_j v_j$, with

$$\alpha_i = (u, v_i), \quad \beta_j = (v, v_j),$$

then

$$\begin{aligned} (u, v) &= (\sum \alpha_i v_i, \sum \beta_j v_j) = \sum \sum \alpha_i \overline{\beta_j} (v_i, v_j) \\ &= \sum \sum \alpha_i \overline{\beta_j} \delta_{ij} = \sum \alpha_i \overline{\beta_i} \\ &= \sum (u, v_i) \overline{(v, v_i)} = \sum (u, v_i) (v_i, v). \end{aligned}$$

(5) \Rightarrow (6). Let $u=v$, then

$$(v, v) = \sum_i \alpha_i \overline{\alpha_i} = \sum_i |\alpha_i|^2 = \sum_i |(v, v_i)|^2$$

(6) \Rightarrow (1). If S is contained in a larger orthonormal set and (6) holds, say if v_0 is orthogonal to each v_i (*i.e.*, v_0 is not in S), then

$$\|v_0\|^2 = \sum |(v_0, v_i)|^2 = \sum 0^2 = 0,$$

so that $v_0=0$. Hence, S is complete.

Q. E. D.

THEOREM 12. (Schwarz's inequality)

If u and v are vectors in an inner product space, then

$$|(u, v)| \leq \|u\| \|v\|$$

PROOF. If $\|u\|=0$ or $\|v\|=0$, then both sides vanish.

If both $\|u\| \neq 0$ and $\|v\| \neq 0$, then

$$\| \alpha u - \beta v \|^2 = (\alpha u - \beta v, \alpha u - \beta v) \geq 0$$

However,

$$\begin{aligned}\|\alpha u - \beta v\|^2 &= \alpha \bar{\alpha}(u, u) - \alpha \bar{\beta}(u, v) - \beta \bar{\alpha}(v, u) + \beta \bar{\beta}(v, v) \\ &= |\alpha|^2(u, u) - \alpha \bar{\beta}(u, v) - \beta \bar{\alpha}(u, v) + |\beta|^2(v, v) \\ &= |\alpha|^2 \|u\|^2 - \alpha \bar{\beta}(u, v) - \bar{\alpha} \beta(\overline{u, v}) + |\beta|^2 \|v\|^2\end{aligned}$$

Now, let us put $\alpha = (v, v) = \|v\|^2$ and $\beta = (u, v)$, then

$$\begin{aligned}\|\alpha u - \beta v\|^2 &= |\alpha|^2 \cdot \|u\|^2 - \alpha \bar{\beta} \cdot \beta - \bar{\alpha} \cdot \beta \cdot \bar{\beta} + |\beta|^2 \cdot \alpha \\ &= |\alpha|^2 \|u\|^2 - \bar{\alpha} |\beta|^2 \\ &= |\alpha|^2 \|u\|^2 - \alpha |\beta|^2 \\ &= \|v\|^4 \|u\|^2 - \|v\|^2 |(u, v)|^2 \geq 0\end{aligned}$$

Dividing both sides by $\|v\|^2$, we have

$$\|v\|^2 \|u\|^2 \geq |(u, v)|^2$$

Q. E. D.

The Schwarz's inequality has important properties:

(1) In any inner product space we define the distance $\delta(u, v)$ between two vectors u and v by

$$\delta(u, v) = \|u - v\| = \sqrt{(u - v, u - v)}$$

In order for δ to deserve distance (metric), it should have the following three properties:

- (1) $\delta(u, v) = \delta(v, u)$,
- (2) $\delta(u, v) \geq 0$; $\delta(u, v) = 0$ iff $u = v$.
- (3) $\delta(u, v) \leq \delta(u, w) + \delta(w, v)$.

Since the above three results are obvious, we omit the proof.

(2) In the n -dimensional Euclidean space R^n , the expression

$$\frac{(u, v)}{\|u\| \|v\|}$$

gives the cosine of the angle between u and v .

The Schwarz's inequality in this case shows that the absolute value of the cosine of a real angle is less than or equal to 1.

(3) In the space F (which is the set of all complex valued continuous functions on the interval $[0, 1]$), the Schwarz inequality becomes

$$\left| \int_0^1 f(t) \overline{g(t)} dt \right|^2 \leq \int_0^1 |f(t)|^2 dt \int_0^1 |g(t)|^2 dt$$

THEOREM 13. If V is an n -dimensional inner product space, then there exist a complete orthonormal set in V , and every complete orthonormal set in V contains exactly n elements.

PROOF. From the corollary 1 of theorem 5, there is a basis $\mathbf{B} = \{u_1, u_2, \dots, u_n\}$ of V , consisting of n independent elements. By the Gram-Schmidt orthogonalization, we have an orthogonal basis

$$\mathbf{B}_0 = \{v_1, \dots, v_n\}$$

of V , and putting

$$\frac{v_i}{\|v_i\|} = e_i,$$

we have a basis

$$\mathbf{B}_0' = \{e_1, \dots, e_n\},$$

which is an orthonormal basis of V .

By the corollary 2 of Theorem 5, and Theorem 11, the rest follows.

Q. E. D.

APPENDIX (GRAM-SCHMIDT ORTHOGONALIZATION)

Let V be an n -dimensional vector space, then there is a basis \mathbf{B} , which has n -independent elements and $\{w_1, \dots, w_m\}$ of \mathbf{B} are orthogonal set:

$$\mathbf{B} = \{w_1, \dots, w_m, v_{m+1}, \dots, v_n\} \quad (n \geq m).$$

If $n=m$, then there is nothing to talk about.

If $n > m$, then let us make a subspace W_{m+1} of V , which is generated by $\{w_1, \dots, w_m, v_{m+1}\}$.

Let us subtract from v_{m+1} its projection among $\{w_1, \dots, w_m\}$. Thus let

$$c_1 = \frac{(v_{m+1}, w_1)}{(w_1, w_1)}, \dots, \quad c_m = \frac{(v_{m+1}, w_m)}{(w_m, w_m)}$$

and let

$$w_{m+1} = v_{m+1} - (c_1 w_1 + \dots + c_m w_m),$$

then (for $1 \leq j \leq m$)

$$\begin{aligned} (w_{m+1}, w_j) &= (v_{m+1} - (\sum c_i w_i), w_j) \\ &= (v_{m+1}, w_j) - \sum_i c_i (w_i, w_j) \\ &= (v_{m+1}, w_j) - c_j (w_j, w_j) \\ &= (v_{m+1}, w_j) - \frac{(v_{m+1}, w_j)}{(w_j, w_j)} (w_j, w_j) \\ &= 0. \end{aligned}$$

Hence, w_{m+1} is perpendicular to $\{w_1, \dots, w_m\}$.

Furthermore, $w_{m+1} \neq 0$, and

$$v_m = w_{m+1} + c_1 w_1 + \dots + c_m w_m;$$

hence v_{m+1} lies in the span of $\{w_1, \dots, w_m, w_{m+1}\}$ (i. e. in W_{m+1}).

Hence $\{w_1, \dots, w_m, w_{m+1}\}$ is an orthogonal basis of W_{m+1} .

We can now proceed by induction, showing that the space W_{n+s} generated by $\{w_1, \dots, w_m, v_{m+1}, \dots, v_{n+s}\}$ has an orthogonal basis

$$\mathbf{B} = \{w_1, \dots, w_m, w_{m+1}, \dots, w_{n+s}\}$$

with $s = 1, 2, \dots, n-m$,

Q. E. D.

Suppose V be a vector space over the field K . Let U and W be subspaces of V . The sum of U and W is the subset of V consisting of all the sums $u+w$ with u in U and w in W .

THEOREM 14. Let U and W be subspaces of V . The sum $U+W$ (the sum of U and W) is a vector space.

PROOF. If u_1, u_2 are in U and w_1, w_2 are in W , then

$$\alpha(u_1+w_1)+\beta(u_2+w_2)=(\alpha u_1+\beta u_2)+(\alpha w_1+\beta w_2)$$

lies in $U+W$, since $\alpha u_1+\beta u_2$ is in U and $\alpha w_1+\beta w_2$ is in W .

Q. E. D.

COROLLARY. Let V be a vector space over K , and let W_i be any subspaces of V , then arbitrary finite sum of W_i is a subspace of V (i. e. $\sum_{i=1}^n W_i$ is a subspace of V).

PROOF. The result is obvious from the Theorem.

Q. E. D.

DEFINITION. We now consider the case of $U+W=V$ and $U \cap W = \{0\}$, then we call that the sum is direct sum of U and W , denoting

$$V = U \oplus W.$$

THEOREM 14. If W is any subspace of a finite-dimensional inner product space V , then V is the direct sum of W and W^\perp , and $W^{\perp\perp} = W$.

PROOF. Let $\mathbf{B} = \{v_1, v_2, \dots, v_n\}$ be an orthonormal set that is complete in W , and let v be any vector in V , we may write $w = \sum_i \alpha_i v_i$, where $\alpha_i = (v, v_i)$. (i. e. w is the projection of v among $\{v_1, \dots, v_n\}$)

It follows from Theorem 10, that $u = v - w$ is in W^\perp , so that v is the sum of two vectors, $v = u + w$, with u in W^\perp and w in W . Hence, $W \cap W^\perp = \{0\}$.

(Because, if v_0 is in $W \cap W^\perp$, then v_0 is in W and v_0 is in W^\perp , hence $(v_0, v_0) = \|v_0\|^2 = 0$).

Therefore, $V = W \oplus W^\perp$.

We observe that in the decomposition $v = u + w$,

We have

$$(v, w) = (u+w, w) = (u, w) + (w, w) = 0 + \|w\|^2 = \|w\|^2,$$

and, similarly,

$$(v, u) = (u+w, u) = (u, u) + (w, u) = \|u\|^2 + 0 = \|u\|^2.$$

Hence, if v is in W^{++} , then $v + u$ (u is in W^+) implies $\|u\|^2 = 0$, so that v is in W (i.e. $W^{++}CW$). And by Lemma 3 WCW^{++} , we have

$$W = W^{++}.$$

Q. E. D

§4. Continuity of a linear map

Let us consider the convergence problems that arise in an inner product space.

DEFINITION. A sequence $\{v_n\}$ of vectors in V converges to a vector v in V if

$$\|v_n - v\| \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty,$$

and we denote $\lim_{n \rightarrow \infty} v_n = v$.

Remark: we shall write N for the dimension of a finite-dimensional vector space, in order to reserve n for the dummy variable in limiting processes.

THEOREM 15. Let V be a finite-dimensional vector space over K . Then the following properties are equivalent:

- (1) $\|v_n - v\| \rightarrow 0$ as $n \rightarrow \infty$
- (2) $(v_n - v, u) \rightarrow 0$ as $n \rightarrow \infty$ for each fixed u in V

PROOF. If (1) is true, then we have for every u , by Schwarz's inequality

$$|(v_n - v, u)| \leq \|v_n - v\| \cdot \|u\| \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty$$

Hence, (1) \Rightarrow (2).

Conversely, assume (2).

Let $\{e_1, e_2, \dots, e_N\}$ be an orthonormal basis in V , then by theorem 11,

$$\|v_n - v\|^2 = \sum_i |(v_n - v, e_i)|^2 \leq N\delta^2,$$

where $\delta^2 = \text{Max}\{|(v_n - v, e_i)|^2\}$

Since δ is arbitrary, $\|v_n - v\| \rightarrow 0$.

Q. E. D.

THEOREM 16. (Cauchy's condition for sequences)

Let V be a finite-dimensional inner product space with the metric defined by the norm, then V is complete (i.e. every Cauchy sequence in V converges).

PROOF. " Let $\{v_n\}$ be a sequence in V such that

$$\|v_m - v_n\| \rightarrow 0 \quad \text{as } m, n \rightarrow \infty "$$

is equivalent to "given $\epsilon > 0$, there is an integer p such that $n > p$ and $m > p$ implies

$$\|v_m - v_n\| < \epsilon "$$

Suppose first that $\lim_{n \rightarrow \infty} v_n = v$, then there is an integer p such that $\|v_n - v\| < \epsilon/2$, for $n > p$.

But we have

$$\|v_m - v_n\| = \|(v_m - v) + (v - v_n)\| \leq \|v_m - v\| + \|v - v_n\|;$$

so that if m and n are both larger than p , it follows that $\|v_m - v\| < \epsilon/2$, $\|v_n - v\| < \epsilon/2$, which implies

$$\|v_m - v_n\| < \epsilon.$$

Since ϵ is arbitrary, $\|v_m - v_n\| \rightarrow 0$.

Conversely, let us suppose the Cauchy condition holds. Let S be the sequence $\{v_n\}$.

If S is finite, then all except a finite number of terms v_n must be equal and $\lim v_n$ will then exist and can be equal to this common value.

If S is infinite, the Cauchy condition implies that S is bounded.

—THE PROOF OF BOUNDEDNESS—

If $\|v_m - v_n\| < \epsilon$ and $M = \text{Max} \{ \|v_1\|, \dots, \|v_n\| \}$, then

$$\|v_m\| = \|(v_m - v_n) + v_n\| \leq \|v_m - v_n\| + \|v_n\|$$

$$< \epsilon + M$$

Since $M + \epsilon$ is finite (because n is fixed), $\{v_n\}$ is bounded,

By the Bolzano-Weierstrass theorem, the sequence must have an accumulation point, say v . Let us assume that the accumulation points be v and v' , then for some $m, n > p$,

$$\|v_m - v\| < \epsilon/3, \|v_n - v\| < \epsilon/3, \|v_m - v_n\| < \epsilon/3$$

and hence

$$\begin{aligned} \|v - v'\| &= \|(v - v_m) + (v_m - v_n) + (v_n - v')\| \\ &\leq \|v - v_m\| + \|v_m - v_n\| + \|v_n - v'\| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon \end{aligned}$$

Hence, the accumulation point is uniquely determined.

Q. E. D.

The metric properties of vectors have certain important implications for the metric properties of linear transformations.

DEFINITION. A linear transformation A on an inner product space V is bounded if there exists a constant M such that $\|Av\| \leq M\|v\|$ for every vector v in V . The greatest lower bound of all constants M with this property is called the norm (or bound) of A and is denoted by $\|A\|$.

Hence, the norm of A is

$$\|A\| = \inf \{M: \|Av\| \leq M\|v\| \text{ for all } v \text{ in } V\}$$

and the notion of boundedness is closely connected with the notion of continuity.

THEOREM 17. Let A be a linear transformation of V and if A be bounded, then A is (uniform) continuous on V .

PROOF. If $\|A\| = 0$, then A must be zero mapping. Hence,

$$\|Au - Av\| = \|A(u-v)\| \leq \|A\| \|u-v\| = 0 < \epsilon \quad (\text{for } \|u-v\| < \delta)$$

implies that A is continuous.

If $\|A\| \neq 0$, then for some positive number ϵ , by writing $\delta = \epsilon / \|A\|$, $\|u-v\| < \delta$ implies

$$\begin{aligned} \|Au - Av\| &= \|A(u-v)\| \leq \|A\| \|u-v\| \leq \|A\| \cdot \delta \\ &\leq \|A\| \cdot \frac{\epsilon}{\|A\|} = \epsilon \end{aligned}$$

Hence, A is (uniformly) continuous on V .

Q. E. D.

THEOREM 18. Every linear transformation on a finite-dimensional inner product space is bounded. (and hence, continuous).

PROOF. Suppose that A is a linear transformation on V and let $\{e_1, \dots, e_N\}$ be an orthonormal basis in V , and write

$$M = \text{Max} \{ \|Ae_1\|, \|Ae_2\|, \dots, \|Ae_N\| \},$$

Since an arbitrary vector v may be written in the form $v = \sum_i (v, e_i) e_i$ by Theorem 11.

We obtain, applying the schwarz inequality and remembering that $\|e_i\| = 1$,

$$\begin{aligned} \|Av\| &= \|A(\sum_i (v, e_i) e_i)\| = \|\sum_i (v, e_i) Ae_i\| \\ &\leq \sum_i |(v, e_i)| \|Ae_i\| \leq \sum_i |(v, e_i)| M \\ &= M \sum_i |(v, e_i)| \leq M \{ \sum_i \|v\| \|e_i\| \} \\ &= M \|v\| N = MN \|v\| \end{aligned}$$

Hence, $\|Av\| \leq MN \|v\|$.

In other words, MN is a bound of A .

Therefore, a linear map A is bounded.

Q. E. D.

Caution : If the underlying vector space V is infinite dimensional, we can not assert that every linear map is bounded.

Conclusions

We have observed various properties of a vector space and basis of a vector space, especially the orthogonal basis.

By some modification, we can always have an orthogonal (and hence, an orthonormal) basis of the vector space.

The norm of a vector is nothing but the length of the vector, however, the norm of a linear transformation is (by definition) just its bound.

The main results of this note are, I suppose,

- (1) We can always have an orthonormal basis in the vector space V .
- (2) Every n -dimensional vector space V over a field K is isomorphic to K^n .
- (3) THEOREM 11, 12, and 14 are very interesting
- (4) The vector space over a field K , with the metric defined by norm is complete, and
- (5) Every bounded linear transformation is (uniformly) continuous on V .

REFERENCES

P. R. Halmos: Finite dimensional vector spaces, Van Nostrand Co., 1958.

S. Lang: Linear Algebra, Addison-Weseley Co., 1966.

T. M. Apostol: Mathematical Analysis, Addison-Weseley Co., 1957.

— 요 지 —

Vector 空間의 基에 關係서

김 영 신

현대 수학에서 vector 공간의 이론은 매우 중요시 되고 있다. 이 논문에서 저자는 다음 사실들을 조사해 보았다.

유한 차원 vector 공간에는 유한 개의 기가 반드시 존재하고 그들을 적당히 변형시킴으로써 직교계를 얻을 수 있으며 내적을 정의함으로써 각과 길이의 개념을 확고히 했으며 길이의 개념은 확장된 개념 (norm) 으로 규정되었고,

- (1) n 차원 vector 공간은 n 차원 좌표 공간과 동형이며
- (2) 유한 차원 vector 공간에서의 계량화를 정리 11, 12, 14 등에서 다루었고
- (3) Vector 공간에 수렴 수열 (Cauchy sequence) 을 도입하여 완비성을 따졌으며
- (4) 선형 변환에 norm을 정의하여 선형 변환과 연속성과의 관계를 따져보았다.