Bibliography:

K. Hoffman and R. Kunze, Linear Algebra.

A. Xepapadeas, Mathematical Methods in Economics.

1 Preliminaries

Note: Small-case **bold** letters represent vectors, i.e. vector $\mathbf{x}$, vector $\mathbf{a}$ etc. but scalar x, scalar a etc.

Let us start by defining the term field.

Definition A set F having at least 2 elements is called a **field** if two operations called addition $(+)$ and multiplication $(\cdot)$ are defined in F and satisfy the following three axioms:

1) For any two elements a, b of F, the sum $a+b$ is defined and is in F; the associative law $(a+b)+c = a+(b+c)$ and the commutative law $a+b = b+a$ hold; and there exists in F a unique element say x such that for any a, b we have that $a+x = b$ (the identity element for this group is denoted by 0 and is called the **zero element** of F).

2) For any two elements a, b of F, the product $ab(= a \cdot b)$ is defined and is in F; the associative law $(ab)c = a(bc)$ and the commutative law $ab = ba$ hold; and there exists in F a unique element say x such that for any a, b with $a \neq 0$, we have that $ax = b$ (the identity element for this group is denoted by 1 and is called the **unit element** or **identity element** of F).

3) The distributive law $a(b+c) = (ab) + (ac) = ab + ac$ holds.

Example of fields are the set of all rational number $\mathbb{Q}$ the set of real numbers $\mathbb{R}$, the set of complex numbers $\mathbb{C}$, etc.

Vector $\mathbf{x}$ with elements $x_1, x_2, ..x_n$ is $n - tuples$ of scalars x_i in F, where F

is a field, i.e. $\mathbf{x} = (x_1, x_2, ..., x_n)$ a raw vector or $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ a column vector. Usually F is the real numbers, denoted by $\mathbb{R}$. In this case we have a **Real Vector**. The set of all real vectors with n elements is denoted by $V_n(\mathbb{R})$.

1.1 Vector Operations

On $V_n(\mathbb{R})$ we say that

$$\mathbf{x} = \mathbf{y} \quad or \quad \mathbf{y} = \mathbf{x} \quad iff \quad x_i = y_i \quad for \quad all \quad i = 1, 2, .., n.$$

For example for $x, y, z \in \mathbb{R}$ we have that $(x + y, x - z, y + z) = (-1, 2, -3)$ iff $x + y = -1$ and $x - z = 2$ $andy + z = -3$, iff $x = 2$, $y = -3$ and $z = 0$.

Furthermore, we say that

$$\mathbf{x} \geq \mathbf{y} \quad (or \quad \mathbf{y} \geq \mathbf{x}) \quad iff \quad x_i \geq y_i \quad (or \quad y_i \geq x_i) \quad for \quad i = 1, 2, .., n.$$

For example if $\mathbf{x} = (-2, 1, 3)$ and $\mathbf{y} = (-3, 0, 3)$ then $\mathbf{x} \geq \mathbf{y}$ as $-2 \geq -3$ and $1 \geq 0$ and $3 \geq 3$. However, notice that can not always say that for any two vectors in $V_n(\mathbb{R})$ with $n > 1$ if there is an inequality relation between them. For example for $\mathbf{z} = (2, 0, -1)$ and $\mathbf{u} = (2, 3, -1)$ then $\mathbf{x} \ngeq \mathbf{y}$ and $\mathbf{x} \nleq \mathbf{y}$, i.e. the two vectors can not be compared.

We can define the multiplication of a vector with a scalar (**scalar multiplication**) by

$$\lambda \mathbf{x} = (\lambda x_1, \lambda x_2, ..., \lambda x_n) \quad for \quad \lambda \in \mathbb{R} \quad and \quad \mathbf{x} \in V_n(\mathbb{R}). \tag{1}$$

The vectors $\mathbf{x}$ and $\lambda \mathbf{x}$ are called **co-linear**. Furthermore, for $\lambda > 0$ we have that if $\mathbf{x} \geq \mathbf{y}$ then $\lambda \mathbf{x} \geq \lambda \mathbf{y}$, whereas if $\lambda < 0$ we have that if $\mathbf{x} \geq \mathbf{y}$ then $\lambda \mathbf{x} \leq \lambda \mathbf{y}$.

The sum of two vectors $\mathbf{x}, \mathbf{y} \in V_n(\mathbb{R})$ is defined as

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) . \tag{2}$$

The commutative and associative Law hold for the summation operation, i.e.

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} \quad and \quad \mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}.$$

In $V_n(\mathbb{R})$ there is the zero element, denoted by $\mathbf{0} = (0, 0, ..., 0)$, and we have that

$$\mathbf{x} + \mathbf{0} = \mathbf{x} \quad for \quad any \quad \mathbf{x} \in \mathbf{V_n}(\mathbb{R}) .$$

2 Vector Spaces

We can now define the vector space:

Definition A **vector space** (or **linear space**) consists of the following:

1. a field F of scalars;

2. a set V of objects, called vectors;

3. a rule (or operation), called vector addition, which associates with each pair of vectors $\mathbf{x}$, $\mathbf{y}$ in V a vector $\mathbf{x} + \mathbf{y}$ in V, called the sum of $\mathbf{x}$ and $\mathbf{y}$, in such a way that

(a) addition is commutative, $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$;

(b) addition is associative, $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$;

(c) there is a unique vector $\mathbf{0}$ in V, called the zero vector, such that $\mathbf{x}+\mathbf{0} = \mathbf{x}$ for all $\mathbf{x}$ in V;

(d) for each vector $\mathbf{x}$ in V there is a unique vector $-\mathbf{x}$ in V such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$;

4. a rule (or operation), called scalar multiplication which associates with each scalar c in F and a vector $\mathbf{x}$ in V a vector $c\mathbf{x}$ in V, called the product of c and $\mathbf{x}$,

in such a way that

(a) $1\mathbf{x} = \mathbf{x}$ for every $\mathbf{x}$ in V;

(b) $(c_1 c_2)\mathbf{x} = c_1(c_2\mathbf{x})$;

(c) $c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$;

(d) $(c_1 + c_2)\mathbf{x} = c_1\mathbf{x} + c_2\mathbf{x}$.

It is important to observe, as the definition states, that a vector space is a composite object consisting of a field, a set of vectors and two operations with certain special properties.

Example 1: The n-tuple space, F^n. Let F be any field, and let V be a set of all n-tuples $\mathbf{x} = (x_1, x_2, ..., x_n)$ of scalars x_i in F. If $\mathbf{y} = (y_1, y_2, ..., y_n)$ with y_i in F, the sum of $\mathbf{x}$ and $\mathbf{y}$ is defined as in (2) and the product of a scalar λ in F and a vector $\mathbf{x}$ in F^n is defined as in (1). The fact that this vector addition and scalar multiplication satisfy condition 3. and 4. is easy to verify, using the similar properties of addition and multiplication of elements of F.

Example 2: The space of functions from a set to a field. Let F be any field and let S be any non-empty set. Let V be the set of all functions from the set S into F. The sum of two vectors f and g in V is the vector $f + g$, i.e. the function from S into F, defined by

$$(f + g)(s) = f(s) + g(s).$$

The product of a scalar λ in F and the function f is the function λf defined by

$$(\lambda f)(s) = \lambda f(s).$$

Again 3. and 4. of the definition are easily verified.

Example 4: The space of polynomial functions over a field F. Let F be a field and let V be the set of all functions f from F into F which have a rule of the form

$$f(x) = a_0 + a_1 x + ... + a_n x^n$$

where $a_0, a_1, ..., a_n$ are fixed scalars in F (independent of the scalar x in F). A function of this type is called a polynomial function on F. Let addition and scalar multiplication be defined as in Example 2. Observe that if f and g are polynomial functions and λ in F, then $f + g$ and λf are again polynomial functions.

Example 5: The space of all $m \times n$ matrices, $F^{m\times n}$ over a field F. Let F be a field and m and n be positive integers. Let $F^{m\times n}$ be the set of all $m \times n$ matrices with elements from F. The sum of two vectors A and B in $F^{m\times n}$ and the product of a scalar a and a vector A are defined in the usual ways. Then $F^{m\times n}$ is a vector space.

A vector $\mathbf{x}$ in V is called a **linear combination** of vectors $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n$ in V if there exist scalars $a_1, a_2, ..., a_n$ in F such that

$$\mathbf{x} = a_1\mathbf{y}_2 + a_2\mathbf{y}_2 + ... + a_n\mathbf{y}_n = \sum_{i=1,..,n} a_i\mathbf{y}_i.$$

It follows that for any $\mathbf{x} \in \mathbb{R}^n$ we have that

$$\mathbf{x} = (x_1, x_2, ..., x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_n + ... + x_n\mathbf{e}_n = \sum_{i=1,..,n} x_i\mathbf{e}_i.$$

Exercises:

1. On $\mathbb{R}^n$ define two operations

$$\begin{aligned} \mathbf{x} \oplus \mathbf{y} &= \mathbf{x} - \mathbf{y} \\ \lambda \otimes \mathbf{x} &= -\lambda\mathbf{x}. \end{aligned}$$

The operations on the right are the usual ones. Which of the axioms for a vector space are satisfied by $(\mathbb{R}^n, \oplus, \otimes)$?

2. Let V be the set of all complex-valued functions f on the real line such that (for all t in $\mathbb{R}$)

$$f(-t) = \overline{f(t)}.$$

The bar denotes complex conjugation. Show that V with the operations

$$\begin{aligned}(f+g)(t) &= f(t)+g(t)\\ (cf)(t) &= cf(t)\end{aligned}$$

is a vector space over the field of real numbers. Give an example of a function in V which is not real-valued.

3 Subspaces

Let V be a vector space over the field F.

Definition A **subspace** of V is a subset W of V which is itself a vectors space over F with the operations of vector addition and scalar multiplication on V.

A direct check of the axioms of the vector space shows that the subset W of V is a subspace if for each $\mathbf{x}$ and $\mathbf{y}$ in W the vector $\mathbf{x}+\mathbf{y}$ is again in W; the $\mathbf{0}$ vector is in W; for each $\mathbf{x}$ in W the vector $(-\mathbf{x})$ is in W; and for each $\mathbf{x}$ in W and each scalar c the vector $c\mathbf{x}$ is in W. The commutativity and associativity of vector addition, and the properties 4. a), b), c), and d) of scalar multiplication do not need to be checked, since these are properties of the operations on V. Things can be simplified still further.

Theorem A non-empty subset W of V is a subspace of V if and only if for each pair of vector $\mathbf{x}$, $\mathbf{y}$ in W and each scalar c in F the vector $c\mathbf{x}+\mathbf{y}$ is again in W.

Proof: Suppose that W is a non-empty subset of V such that $c\mathbf{x}+\mathbf{y}$ belongs to W for all vectors $\mathbf{x}$, $\mathbf{y}$ in W and all scalars c in F. Since W is non-empty, there is a vector $\mathbf{z}$ in W, and consequently $(-1)\mathbf{z} + \mathbf{z} = \mathbf{0}$ is in W. Then if $\mathbf{x}$ is any vector in W and c any scalar in F, the vector $c\mathbf{x} = c\mathbf{x} + \mathbf{0}$ is in W. In particular, $(-1)\mathbf{x} = -\mathbf{x}$ is in W. Finally, if $\mathbf{x}$ and $\mathbf{y}$ in W, then $\mathbf{x} + \mathbf{y} = 1\mathbf{x} + \mathbf{y}$ is in W. Thus W is a subspace of V. Conversely, if W is a subspace of V, $\mathbf{x}$ and $\mathbf{y}$ are in W, and c is a scalar in F, certainly $c\mathbf{x} + \mathbf{y}$ is in W. ■

Examples:

1. If V is any vector space, V is a subspace of V; the subset consisting of the zero vector alone is a subspace of V, called the **zero subspace** of V.

2. If F^n, the set of n-tuples $(x_1, x_2, ..., x_n)$ with $x_1 = 0$ is a subspace; however, the set of n-tuples with $x_1 = 1 + x_2$ is not a subspace $(n \geq 2)$ (why?).

3. The space of polynomial functions over a field F is a subspace of the space of all functions from F into F.

4. The set of all symmetric $n \times n$ matrices is a subspace of the vector space of all $n \times n$ matrices over a field F.

5. The solution space of a system of homogeneous linear equations, i.e. for an $m \times n$ matrix A over a filed F the set of all $(n \times 1)$ vectors $\mathbf{x}$ over F such that $A\mathbf{x} = 0$ (prove that it is a subspace).

Theorem Let V be a vector space over a field F. The intersection of any collection of subspaces of V is a subspace of V.

Let S be a set of vectors in a vector space V. The **subspace spanned** by S is defined to be the the intersection, say W, of all subspaces of V which contain S. When S is a finite set of vectors, $S = \{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_k\}$, we can call W the subspace spanned by the vectors $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_k$.

Theorem The subspace spanned by a non-empty subset S of a vector space V is the set of all linear combinations of vectors in S.

Proof: Let W be the subspace spanned by S. Then each linear combination $\mathbf{y} = a_1\mathbf{s}_1 + a_2\mathbf{s}_2 + ... + a_m\mathbf{s}_m$ of vectors $\mathbf{s}_i$ in S is clearly in W. Thus W contains the set L of all linear combinations of vectors in S, and consequently L is contained in any subspace of V which contains S. Notice also that the set L is non-empty and contains S. If now $\mathbf{x}$ and $\mathbf{y}$ belong to L then $\mathbf{x}$ is a linear combination $\mathbf{x} = b_1\mathbf{s}_1 + b_2\mathbf{s}_2 + ... + b_n\mathbf{s}_n$ of vectors $\mathbf{s}_i$ in S. For each scalar c we have that

$$c\mathbf{x} + \mathbf{y} = cb_1\mathbf{s}_1 + cb_2\mathbf{s}_2 + ... + cb_n\mathbf{s}_n + a_1\mathbf{s}_1 + a_2\mathbf{s}_2 + ... + a_m\mathbf{s}_m$$

which is a linear combination of elements of S and consequently belongs to L. Hence L is a subspace of V. Consequently, L is a subspace which contains S and from above we have that any subspace which contains S contains L. It follows that L is the intersection of all subspaces containing S, i.e. L is the subspace spanned by the set S. ■

If $A_1, A_2, ..., A_k$ are subsets of a vector space V, the set of all sums

$$\mathbf{x}_1 + \mathbf{x}_2 + ... + \mathbf{x}_k$$

of vectors $\mathbf{x}_i$ in A_i is called the **sum** of the subsets $A_1, A_2, ..., A_k$ and is denoted by

$$A_1 + A_2 + ... + A_k.$$

If $W_1, W_2, ..., W_k$ are subspaces of a vector space V then the sum

$$W = W_1 + W_2 + ... + W_k$$

is easily seen to be a subspace of V which contains each of the subspaces W_i. From this follows that W is the subspace spanned by the union of $W_1, W_2, ..., W_k$.

Example: Let F be a subfield of the field C of complex numbers. Assume that $\mathbf{s}_1 = (1, 2, 0, 3, 0)$, $\mathbf{s}_2 = (0, 0, 1, 1, 0)$, and $\mathbf{s}_3 = (0, 0, 0, 0, -1)$. A vector $\mathbf{x}$ is in the subspace W of F^5 spanned by $\mathbf{s}_1$, $\mathbf{s}_2$, $\mathbf{s}_3$ if and only if there exist scalars a_1,

a_2 and a_3 in F such that

$$\begin{aligned}\mathbf{x} &= a_1\mathbf{s}_1 + a_2\mathbf{s}_2 + a_3\mathbf{s}_3 \\ &= (a_1, 2a_1, 0, 3a_1, 0) + (0, 0, a_2, a_2, 0) + (0, 0, 0, 0, -a_3) \\ &= (a_1, 2a_1, a_2, 3a_1 + a_2, -a_3)\,.\end{aligned}$$

Hence W consists of all vectors of the form where a_1, a_2 and a_3 arbitrary scalars in F. Alternatively, W can be described as the set of all 5-tuples $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$ with x_i in F such that

$$\begin{aligned}x_2 &= 2x_1 \\ x_4 &= 3x_1 + x_3.\end{aligned}$$

Thus the vector $(-1, -2, 5, 2, -7)$ is in W, whereas $(1, -2, 5, 2, -7)$ is not.

Example: Let V be the space of all polynomial functions over F. Let S be the subspace of V consisting of the polynomial functions f_0, f_1, f_2,... defined by

$$fn\,(x) = x^n \quad where \quad n = 0, 1, 2, .. \quad and \quad x \in F.$$

Then V is the subspace spanned by the set S.

Example: Let F be a subfield of the field $\mathbb{C}$ of complex numbers, and let V be the vector space of all 2x2 matrices over F. Let W_1 be the subset of V consisting of all matrices of the form $\begin{pmatrix} x & y \\ z & 0 \end{pmatrix}$ where x, y, z are arbitrary scalars in F. Finally, let W_2 be the subset of V consisting of all matrices of the form $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ where a, b are arbitrary scalars in F. Then W_1 and W_2 are subspaces of V. Also $V = W_1 + W_2$ because for any matrix $\begin{pmatrix} a & b \\ d & f \end{pmatrix}$ in V we have that

$$\begin{pmatrix} a & b \\ d & f \end{pmatrix} = \begin{pmatrix} a & b \\ d & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & f \end{pmatrix}.$$

The subspace $W_1 \cap W_2$ consists of all matrices of the form $\begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix}$ where x are arbitrary scalars in F.

Exercises:

1. Which of the following sets of vectors $\mathbf{x} = (x_1, x_2, ..., x_n)'$ in $\mathbb{R}^n$ are subspaces of $\mathbb{R}^n$ $(n \geq 3)$?

a) all $\mathbf{x}$ such that $x_1 \geq 0$;

b) all $\mathbf{x}$ such that $x_1 + 3x_2 = x_3$;

c) all $\mathbf{x}$ such that $x_2 = x_1^2$;

d) all $\mathbf{x}$ such that $x_1 x_2 = 0$;

e) all $\mathbf{x}$ such that x_1 is rational.

2. Let V be the (real) vector space of all functions f from $\mathbb{R}$ into R. Which of the following sets of functions are subspaces of V?

a) all f such that $f(x^2) = (f(x))^2$;

b) all f such that $f(0) = f(1)$;

c) all f such that $f(3) = 1 + f(-5)$;

d) all f such that $f(-1) = f(0)$;

e) all f which are continuous.

3. Let F be a field and let n be a positive integer $(n \geq 2)$. Let V be the vector space of all $n \times n$ matrices over F. Which of the following sets of matrices A in V are subspaces of V?

a) all invertible A;

b) all non-invertible A;

c) all A such that $AB = BA$, where B is some fixed matrix in V;

d) all A such that $A^2 = A$.

4. Let W_1 and W_2 be subspaces of a vector space V such that the union of W_1 and W_2 is also a subspace. Prove that one of the spaces W_1 or W_2 is contained in

the other.

4 Bases and Dimensions

Definition Let V be a vector space over F. A subset S of V is said to **linearly dependent** if there exist distinct vectors $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n$ is S and scalars a_1, a_2,..,a_n in F not all of which are 0, such that

$$a_1\mathbf{s}_1 + a_2\mathbf{s}_2 + ... + a_n\mathbf{s}_n = \mathbf{0}.$$

A set which is not linearly dependent is called **linearly independent**. If the set S contains only finitely many vectors $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n$, we say that $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n$ are independent (or dependent) instead of saying that S is independent (or dependent).

It is easy to prove that:

1. Any set which contains a linearly dependent set is linearly dependent.

2. Any subset of a linearly independent set is linearly independent.

3. A set S of vectors is linearly independent if and only if each finite subset of S is linearly independent, i.e. if and only if for any distinct vectors $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_n$ of S, and $a_1\mathbf{s}_1 + a_2\mathbf{s}_2 + ... + a_n\mathbf{s}_n = \mathbf{0}$ implies each $a_i = 0$.

4. Any set which contains the $\mathbf{0}$ vector is linearly dependent.

Let V be a vector space.

Definition A **basis** for V is a linearly independent set of vectors in V which spans the space V. The space V is **finite-dimensional** if it has a finite basis.

Examples:

1. Let F be any field and let S be the subset of F^n consisting of the vectors

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

These vectors have all their elements 0 apart from one which equals 1 and are called **elementary vectors** $\mathbf{e}_j$ $j = 1, ..., n$. Hence the $\mathbf{e}_j$ vector is

$$\mathbf{e}_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Now let $x_1, x_2, ..., x_n$ be scalars in F and put $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n$. Then

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

This shows that $\mathbf{e}_1, \mathbf{e}_n, ..., \mathbf{e}_n$ span F^n. Since $\mathbf{x} = \mathbf{0}$ if and only if $x_1 = x_2 = ... = x_n = 0$, the vectors $\mathbf{e}_1, \mathbf{e}_n, ..., \mathbf{e}_n$ are linearly independent. The set $S = \{\mathbf{e}_1, \mathbf{e}_n, ..., \mathbf{e}_n\}$ is accordingly a basis for F^n, called the **standard basis** of F^n. Furthermore, for any vector $\mathbf{x}$ in V the scalars x_i $i = 1, ..., n$ are called the **standard basis coordinates** of $\mathbf{x}$ or simply **coordinates**.

2. Let P be an invertible $n \times n$ matrix with entries in a field F. Then $P_1, ...P_n$, the columns of P, form a basis for the space of column matrices (vectors) $F^{n\times1}$. To

prove this notice that if $\mathbf{x}$ is a column matrix (vector) then $P\mathbf{x} = x_1 P_1 + ... + x_n P_n$. Since $P\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$, it follows that $\{P_1, ...P_n\}$ is a linearly independent set. Now to prove that it spans $F^{n\times 1}$ consider any vector $\mathbf{y}$ in $F^{n\times 1}$. Then if $\mathbf{x} = P^{-1}\mathbf{y}$ we have that $\mathbf{y} = P\mathbf{x}$ which means that $\mathbf{y} = x_1 P_1 + ... + x_n P_n$. Hence $\{P_1, ..., P_n\}$ is a basis in $F^{n\times 1}$.

3. This is an example of an infinite basis. Let F be a subfield of the complex numbers and let V be the space of polynomial functions over F, i.e. functions from F to F which have a rule of the form

$$f(x) = c_0 + c_1 x + ... + c_n x^n.$$

Let $f_k(x) = x^k$, $k = 0, 1, 2, ...$ The infinite set $\{f_0, f_1, f_2, ...\}$ is a basis for V. Clearly the set spans V, as the function above f is

$$f(x) = c_0 f_0 + c_1 f_1 + ... + c_n f_n.$$

Hence we have to prove that set $\{f_0, f_1, f_2, ...\}$ is independent. It suffices to show that each finite subset of it is independent, i.e. for each n the set $\{f_0, f_1, f_2, ..., f_n\}$ is independent. Suppose that

$$c_0 f_0 + c_1 f_1 + ... + c_n f_n = 0$$

i.e.

$$c_0 + c_1 x + ... + c_n x^n = 0$$

for every x in F. This is to say that every x in F is a root of the polynomial $f(x) = c_0 + c_1 x + ... + c_n x^n$. However, a polynomial of degree n in $\mathbb{C}$ has only n roots in $\mathbb{C}$. Hence $c_0 = c_1 = ... = c_n = 0$.

We have given an infinite basis for V. Does this mean that V is not finite-dimensional? This is true if we prove that V can not have a finite basis. Suppose we have a finite number of polynomial functions $g_1, g_2, ..., g_k$. There will be a largest power of x which has non-zero coefficient in $g_1(x), g_2(x), ..., g_k(x)$. If that power is m clearly $f_{m+1}(x) = x^{m+1}$ is not in the linear span of $g_1, g_2, ..., g_k$. So V is not finite-dimensional.

Theorem Let V be a vector space which is spanned by a finite set of vectors $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_m$. Then any independent set of vectors in V is finite and contains no more than m elements.

Proof: To prove the theorem it suffices to show that every subset W in V which contains more than m vectors is linearly dependent. Let W be such a set with elements $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n$ where $n > m$. Since $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_m$ span V, there exist scalars a_{ij} in F such that

$$\mathbf{w}_j = \sum_{i=1}^{m} a_{ij}\mathbf{s}_i.$$

For any n scalars $b_1, b_2, ..., b_n$ we have that

$$\begin{aligned} b_1\mathbf{w}_1 + b_2\mathbf{w}_2 + ... + b_n\mathbf{w}_n &= \sum_{j=1}^{n} b_j\mathbf{w}_j = \sum_{j=1}^{n} b_j \sum_{i=1}^{m} a_{ij}\mathbf{s}_i \\ &= \sum_{j=1}^{n}\sum_{i=1}^{m} (b_j a_{ij})\,\mathbf{s}_i \\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} b_j a_{ij}\right)\mathbf{s}_i, \end{aligned}$$

and calling $\sum_{j=1}^{n} b_j a_{ij} = c_i$. Hence, if $b_1\mathbf{w}_1 + b_2\mathbf{w}_2 + ... + b_n\mathbf{w}_n = 0$ we get that $\sum_{i=1}^{m} c_i\mathbf{s}_i = 0$. As now $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_m$ span V, we have that $c_1 = c_2 = ... = c_m = 0$. Hence

$$\sum_{j=1}^{n} b_j a_{1j} = 0, \quad \sum_{j=1}^{n} b_j a_{2j} = 0, \quad ..., \quad \sum_{j=1}^{n} b_j a_{nj} = 0$$

and it follows

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \mathbf{0}.$$

This is a homogeneous system of m equations and n unknowns with $m < n$. Consequently, there are at least $n - m$ non-zero $b_i\prime s$ which satisfy the system. This in fact means for $b_1\mathbf{w}_1 + b_2\mathbf{w}_2 + ... + b_n\mathbf{w}_n = 0$ not all $b_i\prime s$, for $i = 1, ..., n$, should be 0. Hence, W is a linearly dependent set. ■

Corollary If V is a finite-dimensional vector space, then any two bases of V have the same finite number of elements. Then any independent set of vectors in V is finite and contains no more than m elements.

Proof: Since V is a finite-dimensional vector space it has a finite basis $\{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_m\}$. From the above theorem every basis of V is finite and contains no more than m elements. Hence if $\{\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n\}$ is a basis then $n \leq m$. But if $\{\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n\}$ is a basis then $m \leq n$. Hence $m = n$. ■

The above corollary allows us to define the **dimension** of a finite-dimensional vector space as the number of elements in a basis for V. This allows us to reformulate the above theorem as follows:

Corollary Let V be a finite-dimensional vector space and let $n = \dim(V)$. Then

(1) any subset of V which contains more than n vectors is linearly dependent;

(2) no subset of V contains fewer than n vectors can span V.

Examples: If F is a field, the dimension F^n is n, because the standard basis for F^n contains n vectors. The matrix space $F^{m\times n}$ has dimension mn, because the matrices which have a 1 in the i, j place and 0 elsewhere form a basis for $F^{m\times n}$.

If V is any vector space over F, the zero subspace of V is spanned by the vector $\mathbf{0}$, but $\{\mathbf{0}\}$ is a linearly dependent set and consequently not a basis. For this reason we say that the zero subspace has dimension 0.

Lemma Let S be a linearly independent subset of a vector space V. Suppose that $\mathbf{x}$ is a vector in V which is not in the subspace spanned by S. Then the set obtained by adjoining $\mathbf{x}$ to S is linearly independent.

Theorem If W is a subspace of a finite-dimensional vector space V, every linearly independent subset of W is finite and is part of a finite basis for W.

Corollary If W is a proper subspace of a finite-dimensional vector V, then W is finite-dimensional and $\dim(W) < \dim(V)$.

Theorem If W_1 and W_2 are finite-dimensional subspaces of a vector space V, then $W_1 + W_2$ is finite-dimensional and

$$\dim(W_1) + \dim(W_2) = \dim(W_1 \cap W_2) + \dim(W_1 + W_2).$$

Exercises:

1. Let V be a vector space over a subfield F of the complex numbers. Suppose $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ are linearly independent vectors in V. Prove that $(\mathbf{x}+\mathbf{y})$, $(\mathbf{y}+\mathbf{z})$, and $(\mathbf{z}+\mathbf{x})$ are linearly independent.

2. Let V be a vector space over the field F. Suppose there are finite number of vectors $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ in V which span V. Prove that V is finite-dimensional.

5 Coordinates

For this subsection all vectors are column vectors. One of the useful features of a basis $\mathcal{B}$ in an n-dimensional space V is that it essentially enables one to introduce coordinates in V analogous to the 'natural coordinates' x_i of a vector $\mathbf{x} = (x_1, ..., x_n)'$ in the space F^n. In this respect the coordinates of a vector $\mathbf{x}$ in V relative to the basis $\mathcal{B}$ will be the scalars which serve to express $\mathbf{x}$ as a linear combination of the vectors in the basis. If $\mathcal{B}$ is an arbitrary basis of the n-dimensional vector space V, we shall probably have no natural ordering of the vectors in $\mathcal{B}$. Consequently, we have to assume an ordering of the vectors in $\mathcal{B}$ and we shall refer to this basis as **ordered basis**. In such a way, for and ordered basis $\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_n\}$, for any vector $\mathbf{x}$ in V there is a unique n-tuple $(x_1, ..., x_n)$ of scalars in F such that $\mathbf{x} = \sum_{i=1}^{n} x_i b_i$. Hence, x_i is the i^{th} **coordinate** of $\mathbf{x}$ **relative to the ordered basis** $\mathcal{B}$. Consequently, we can say that each ordered basis for V determines a one-one correspondence

$$\mathbf{x} \to (x_1, ..., x_n)^{/}$$

between the set of all vectors in V and the set of all n-tuples in F^n. This correspondence has the property that the correspondent of $\mathbf{x} + \mathbf{y}$ is the sum in F^n of the correspondences of $\mathbf{x}$ and $\mathbf{y}$, and that the corresponded of $c\mathbf{x}$ is the product in F^n of the scalar c and the correspondent of $\mathbf{x}$. To indicate the dependence of the coordinates on the basis the symbol $[\mathbf{x}]_{\mathcal{B}}$ is used for the coordinates of $\mathbf{x}$ relative to the ordered basis $\mathcal{B}$.

Theorem Let V be a n-dimensional vector space over the field F, and let $\mathcal{B}$ and $\mathcal{B}^{/}$ be two ordered bases of V. Then there is a unique invertible $n \times n$ matrix P with entries in F such that

$$[\mathbf{x}]_{\mathcal{B}} = P\,[\mathbf{x}]_{\mathcal{B}^{/}} \quad \textit{and} \quad [\mathbf{x}]_{\mathcal{B}^{/}} = P^{-1}\,[\mathbf{x}]_{\mathcal{B}}$$

for any vector $\mathbf{x}$ in V.

Theorem Let P is an $n \times n$ invertible matrix over F. Let V be an n-dimensional vector space over the field F, and let $\mathcal{B}$ be an ordered basis of V. Then there is a unique ordered basis $\mathcal{B}^{/}$ of V such that

$$[\mathbf{x}]_{\mathcal{B}} = P\,[\mathbf{x}]_{\mathcal{B}^{/}} \quad \textit{and} \quad [\mathbf{x}]_{\mathcal{B}^{/}} = P^{-1}\,[\mathbf{x}]_{\mathcal{B}}$$

for any vector $\mathbf{x}$ in V.

Example: Let F be a field and let $\mathcal{B}$ be the standard ordered basis of F^n. Let $\mathbf{x} = (x_1, ..., x_n)^{/}$ be a vector in F^n. Then $[\mathbf{x}]_{\mathcal{B}} = (x_1, ..., x_n)^{/}$.

Example: Let $\mathbb{R}$ be the field of real numbers and let θ be a fixed real number. The matrix and

$$P = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

is invertible with inverse

$$P^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

Thus for each θ the set $\mathcal{B}^{/}$ consisting of the vectors $(\cos\theta, \sin\theta)^{/}$, $(-\sin\theta, \cos\theta)^{/}$ is a basis for $\mathbb{R}^2$. Intuitively this basis may be described as the one obtained by rotating the standard basis through the angle θ. If $\mathbf{x} = (x_1, x_2)^{/}$ is a vector in $\mathbb{R}^2$, then

$$[\mathbf{x}]_{\mathcal{B}^{/}} = P^{-1}[\mathbf{x}]_{\mathcal{B}} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1\cos\theta + x_2\sin\theta \\ -x_1\sin\theta + x_2\cos\theta \end{pmatrix},$$

or the coordinates relative to $\mathcal{B}^{/}$ are

$$\begin{pmatrix} x_1^{/} \\ x_2^{/} \end{pmatrix} = \begin{pmatrix} x_1\cos\theta + x_2\sin\theta \\ -x_1\sin\theta + x_2\cos\theta \end{pmatrix}.$$

Example: Let F be a subfield of the complex numbers. The matrix

$$P = \begin{pmatrix} -1 & 4 & 5 \\ 0 & 2 & -3 \\ 0 & 0 & 8 \end{pmatrix}$$

is invertible with inverse

$$P^{-1} = \begin{pmatrix} -1 & 2 & \frac{11}{8} \\ 0 & \frac{1}{2} & \frac{3}{16} \\ 0 & 0 & \frac{1}{8} \end{pmatrix}.$$

Thus the vectors $(-1,0,0)^{/}$, $(4,2,0)^{/}$ and $(5,-3,8)^{/}$ form a basis $\mathcal{B}^{/}$ of F^3. The coordinates $x_1^{/}$, $x_2^{/}$, $x_3^{/}$ of the vector $\mathbf{x} = (x_1, x_2, x_3)^{/}$ relative to the basis $\mathcal{B}^{/}$ are given by

$$\begin{pmatrix} x_1^{/} \\ x_2^{/} \\ x_3^{/} \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 & 2 & \frac{11}{8} \\ 0 & \frac{1}{2} & \frac{3}{16} \\ 0 & 0 & \frac{1}{8} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_1 + 2x_2 + \frac{11}{8}x_3 \\ \frac{1}{2}x_2 + \frac{3}{16}x_3 \\ \frac{1}{8}x_3 \end{pmatrix}.$$

Exercises:

1. Show that the vectors

$$\begin{aligned} \mathbf{s}_1 &= (1,1,0,0)^{/}, \quad \mathbf{s}_2 = (0,0,1,1)^{/} \\ \mathbf{s}_1 &= (1,0,0,4)^{/}, \quad \mathbf{s}_4 = (0,0,0,2)^{/} \end{aligned}$$

form a basis for $\mathbb{R}^4$. Find the coordinates of each of the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ with respect to the ordered basis $\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4\}$.

2. Let V be the vector space over the complex numbers of all functions from $\mathbb{R}$ into $\mathbb{C}$, i.e. the space of all complex-valued functions on the real line. Let $f_1(x) = 1$, $f_2(x) = e^{ix}$, and $f_3(x) = e^{-ix}$. a) Prove that $f_1(x) = 1$, $f_2(x) = e^{ix}$, and $f_3(x) = e^{-ix}$ are linearly independent. b) Let $g_1(x) = 1$, $g_2(x) = \cos x$, and $g_3(x) = \sin x$. Find an invertible 3×3 matrix P such that

$$g_i = \sum_{j=1}^{3} P_{ij} f_j \quad for \quad i = 1, 2, 3.$$

6 Inner Product Spaces

In this subsection we consider only the cases that F is a subfield of either $\mathbb{R}$ the field of real numbers or $\mathbb{C}$ the field of complex numbers.

Let F be the field of real or complex numbers, and V a vector space over F.

Definition An **inner product** on V is a function which assigns to each ordered pair of vectors $\mathbf{x}$, $\mathbf{y}$ in V a scalar $(\mathbf{x} \bullet \mathbf{y})$ in F in such a way that for all $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ in V and all scalars c in F

(a) $(\mathbf{x}+\mathbf{y}) \bullet \mathbf{z} = \mathbf{x} \bullet \mathbf{z} + \mathbf{y} \bullet \mathbf{z}$

(b) $(c\mathbf{x} \bullet \mathbf{y}) = c\,(\mathbf{x} \bullet \mathbf{y})$

(c) $\mathbf{x} \bullet \mathbf{y} = \overline{\mathbf{y}} \bullet \overline{\mathbf{x}}$ where the bar denotes complex conjugation

(d) $\mathbf{x} \bullet \mathbf{x} > 0 \quad if \quad \mathbf{x} \neq \mathbf{0}$.

It should be observed that conditions (a), (b), and (c) imply

(e) $\mathbf{x} \bullet (c\mathbf{y} + \mathbf{z}) = \overline{c}\mathbf{x} \bullet \mathbf{y} + \mathbf{x} \bullet \mathbf{z}$.

When F is the field $\mathbb{R}$ of real numbers then for vector $\mathbf{x}$ in $\mathbb{R}$ we have that $\overline{\mathbf{x}} = \mathbf{x}$.

Example 1: On F^n there is an inner product which we call the **standard inner product**. For any two vectors $\mathbf{x} = (x_1, x_2, ..., x_n)$, $\mathbf{y} = (y_1, y_2, ..., y_n)$ in F^n, the standard inner product is defined as the scalar:

$$\mathbf{x} \bullet \mathbf{y} = x_1\overline{y_1} + x_2\overline{y_2} + ... + x_n\overline{y_n} = \sum_{i=1,..,n} x_i\overline{y_i}.$$

Specifically if $F = \mathbb{R}$ we have

$$\mathbf{x} \bullet \mathbf{y} = x_1y_1 + x_2y_2 + ... + x_ny_n = \sum_{i=1,..,n} x_iy_i.$$

Furthermore, if $\mathbf{x}, \mathbf{y}$ are column vectors then

$$\mathbf{x} \bullet \mathbf{y} = \mathbf{x}^{/}\mathbf{y} = x_1y_1 + x_2y_2 + ... + x_ny_n = \sum_{i=1,..,n} x_iy_i.$$

In such a way the standard inner product agrees with the general rule of matrix multiplication, i.e. $1 \times n$ matrix times $n \times 1$ matrix is a scalar (1×1).

Example 2: Consider a consumer who purchases n products at prices $\mathbf{p} = (p_1, p_2, ..., p_n)$ and at quantities $\mathbf{q} = (q_1, q_2, ..., q_n)$. Then his total expenditure is given by the inner product of $\mathbf{p}$ and $\mathbf{q}$, i.e. the *total expenditure* $= \mathbf{p} \bullet \mathbf{q} = p_1q_1 + p_2q_2 + ... + p_nq_n = \sum_{i=1,..,n} p_iq_i$.

Example 3: Let V be the vector space of all continuous complex-valued functions on the unit interval, $0 \leq t \leq 1$. Let

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

If f, g are real valued functions on the unit interval, then the complex conjugate on g may be omited and we can define the inner product for this space as

$$f \bullet g = \int_0^1 f(t)\,g(t)\,dt.$$

Example 4: One way of constructing new inner products from a given one is the following. Let V and W be vector spaces over F and suppose $(\bullet)$ is an inner product on W. If T is a non-singular linear transformation from V into W, then the equation

$$p_T(\mathbf{x}, \mathbf{y}) = (T\mathbf{x} \bullet T\mathbf{y})$$

defines an inner product p_T on V. The following are a special cases.

(1) Let V be a finite dimensional vector space, and let $\mathcal{B} = \{\mathbf{b}_1, ..., \mathbf{b}_n\}$ be an ordered basis for V. Let $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ be the standard basis vector inF^n, and let T be the linear transformation from V into F^n such that $T\mathbf{b}_j = \mathbf{e}_j$, $j = 1, ..., n$. T is called the natural isomorphism of V onto F^n, then

$$p_T\left(\sum_{j=1}^n x_j\mathbf{b}_j, \sum_{i=1}^n y_i\mathbf{b}_i\right) = \sum_{j=1}^n x_j\overline{y_i}.$$

Thus for any basis for V there is an inner product on V with the property $(\mathbf{b}_j \bullet \mathbf{b}_i) = \delta_{ji}$. In fact, it is to show that there is exactly one such inner product and it is determined by the basis $\mathcal{B}$ for V.

(2) Let V be the vector space of all continuous complex-valued functions on the unit interval, $0 \leq t \leq 1$ (see example 3 above). Let T be the linear operator

'multiplication by t', that is $(Tf)\,t = tf\,(t)$, $0 \leq t \leq 1$. It is easy to see that T is linear. Also T is non-singular; for suppose that $Tf = 0$. Then $tf(t) = 0$ for all $0 \leq t \leq 1$; hence $f(t) = 0$ for $t > 0$. Since f is continuous, we have $f(0) = 0$ as well and consequently $f = 0$. Now using the inner product of the example 3 we construct a new inner product on V by setting

$$\begin{aligned} p_T\,(f, g) &= \int_0^1 (Tf)\,(t)\,\overline{(Tg)\,(t)}dt \\ &= \int_0^1 t^2 f\,(t)\,\overline{g\,(t)}dt. \end{aligned}$$

Occasionally it is very useful to know that an near product on a real or complex vector space is determined by another function, the so-called quadratic form determined by the inner product. To define it, we first denote the positive square root of $(\mathbf{x} \bullet \mathbf{x})$ by $\|\mathbf{x}\|$ and it is called the **norm** of $\mathbf{x}$ with respect to the inner product. For example in $\mathbb{R}^3$ and for the standard inner product we have that

$$\|\mathbf{x}\| = (\mathbf{x} \bullet \mathbf{x})^{\frac{1}{2}} = \left(x_1^2 + x_2^2 + x_3^2\right)^{\frac{1}{2}}$$

i.e. the length of $\mathbf{x}$. The **quadratic form** determined by the inner product is the function that assigns to each vector $\mathbf{x}$ the scalar $\|\mathbf{x}\|^2$.

Now we can define the notion of inner product space. An i**nner product space** is a real or complex vector space, together with a specified inner product on that space. A complex inner product space is often referred to as a **unitary space**. A finite-dimensional real inner product space is often called a **Euclidean space**.

Theorem If V is an inner product space, then for any vectors $\mathbf{x}$, $\mathbf{y}$ in V and any scalar c

(i) $\|c\mathbf{x}\| = |c|\,\|\mathbf{x}\|$

(ii) $\|\mathbf{x}\| > 0 \quad for \quad \mathbf{x} \neq \mathbf{0}$

(iii) $\|\mathbf{xy}\| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$

(iv) $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Proof: (i) and (ii) follow almost immediately from the various definitions involved. The inequality in (iii) is clearly valid when $\mathbf{x} = \mathbf{0}$. If $\mathbf{x} \neq \mathbf{0}$, put

$$\mathbf{z} = \mathbf{y} - \frac{(\mathbf{y} \bullet \mathbf{x})}{\|\mathbf{x}\|^2}\mathbf{x}.$$

Then $(\mathbf{z} \bullet \mathbf{x}) = 0$ and

$$\begin{aligned} 0 \;\leq\; \|\mathbf{z}\|^2 &= \left(\left(\mathbf{y} - \frac{(\mathbf{y} \bullet \mathbf{x})}{\|\mathbf{x}\|^2}\mathbf{x}\right) \bullet \left(\mathbf{y} - \frac{(\mathbf{y} \bullet \mathbf{x})}{\|\mathbf{x}\|^2}\mathbf{x}\right)\right) \\ &= \mathbf{y} \bullet \mathbf{y} - \frac{(\mathbf{y} \bullet \mathbf{x})\,(\mathbf{x} \bullet \mathbf{y})}{\|\mathbf{x}\|^2} = \|\mathbf{y}\|^2 - \frac{|(\mathbf{x} \bullet \mathbf{y})|^2}{\|\mathbf{x}\|^2}. \end{aligned}$$

Hence $|(\mathbf{x} \bullet \mathbf{y})|^2 \leq \|\mathbf{x}\|^2 \, \|\mathbf{y}\|^2$. Now

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\|^2 &= (\mathbf{x}+\mathbf{y}) \bullet (\mathbf{x}+\mathbf{y}) = (\mathbf{x} \bullet \mathbf{x}) + (\mathbf{y} \bullet \mathbf{x}) + (\mathbf{x} \bullet \mathbf{y}) + (\mathbf{y} \bullet \mathbf{y}) \\ &= \|\mathbf{x}\|^2 + \overline{(\mathbf{x} \bullet \mathbf{y})} + (\mathbf{x} \bullet \mathbf{y}) + \|\mathbf{y}\|^2 \end{aligned}$$

as $\overline{(\mathbf{x} \bullet \mathbf{y})} = \mathbf{y} \bullet \mathbf{x}$ (see (c) above). Hence

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \mathbf{2}\,\mathrm{Re}\,(\mathbf{x} \bullet \mathbf{y}) + \|\mathbf{y}\|^2$$

where $\mathrm{Re}\,(z)$ is the real part of a complex number z. As now from (iii) we have that $\|(\mathbf{x} \bullet \mathbf{y})\| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$ it follows that $\mathrm{Re}\,(\mathbf{x} \bullet \mathbf{y}) \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$. Hence we have that

$$\|\mathbf{x}+\mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + \mathbf{2}\,\|\mathbf{x}\| \, \|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2$$

and it follows that

$$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$$

■

The inequality in (iii) is called the **Cauchy-Schwarz inequality** and has a wide variety of applications. The proof shows that if $\mathbf{x}$ is non-zero then $|(\mathbf{x} \bullet \mathbf{y})| <$

$\|\mathbf{x}\| \|\mathbf{y}\|$ unless $\mathbf{y} = \frac{(\mathbf{y} \bullet \mathbf{x})}{\|\mathbf{x}\|^2}\mathbf{x}$. Hence equality occurs in (iii) if and only if $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent.

Example 5: Applying the Cauchy-Schwarz inequality to the inner product in Example 1 we get

$$\sum_{i=1,..,n} x_i \overline{y_i} \leq \left(\sum_{i=1,..,n} |x_i|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1,..,n} |y_i|^2 \right)^{\frac{1}{2}}$$

and in Example 3 we get

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

Definition If $\mathbf{x}$ and $\mathbf{y}$ are vectors in an inner product space V and $(\mathbf{x} \bullet \mathbf{y}) = 0$ then $\mathbf{x}$ is called **orthogonal** to $\mathbf{y}$. Since this implies that $\mathbf{y}$ is orthogonal to $\mathbf{x}$ we simply say that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal. If S is a set of vectors in V, S is called an **orthogonal set** provided that all pairs of distinct vectors in S are orthogonal. An orthonormal set is an orthogonal set S with the additional property that $\|\mathbf{x}\| = 1$, for every $\mathbf{x}$ in S.

The zero vector is orthogonal to any vector in V and it is the only vector with this property.

Example 6: The standard basis of either $\mathbb{R}^n$ or $\mathbb{C}^n$ is an orthonormal set with respect to the standard inner product.

Example 7: The vector $(x, y)^{/}$ in $\mathbb{R}^2$ is orthogonal to $(-y, x)^{/}$ with respect to the standard inner product as

$$(x, y) \begin{pmatrix} -y \\ x \end{pmatrix} = -xy + xy = 0.$$

Example 8: Let V be the space of continuous complex valued (or real valued) functions on the interval $0 \leq x \leq 1$ with the inner product

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

Suppose that $f_n(x) = \sqrt{2}\cos(2\pi nx)$ and that $g_n(x) = \sqrt{2}\sin(2\pi nx)$. Then $\{1, f_1, g_1, f_2, g_2, ...\}$ is an infinite orthonormal set. We may also form the linear combinations

$$\frac{1}{\sqrt{2}}(f_n + ig_n) \quad n = 1, 2, 3, ...$$

In this way we get a new orthonormal which consists of all functions of the form

$$\frac{1}{\sqrt{2}}h_n = \exp(2\pi inx) \quad n = \pm 1, \pm 2, \pm 3, ...$$

Theorem An orthogonal set of non-zero vectors is linearly independent.

Proof: Let S be a finite or infinite orthogonal set of non-zero vectors in a given inner product space. Suppose that $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ are distinct vectors in S and that

$$\mathbf{y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + ... + a_k\mathbf{x}_k.$$

Then

$$(\mathbf{y} \bullet \mathbf{x}_i) = ((a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + ... + a_k\mathbf{x}_k) \bullet \mathbf{x}_i) = a_i(\mathbf{x}_i \bullet \mathbf{x}_i)$$

as $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ orthogonal and as $(\mathbf{x}_i \bullet \mathbf{x}_i) \neq 0$ it follows that

$$a_i = \frac{(\mathbf{y} \bullet \mathbf{x}_i)}{(\mathbf{x}_i \bullet \mathbf{x}_i)} \quad for \quad i = 1, ..., k.$$

Hence if $\mathbf{y} = \mathbf{0}$ we have that each $a_i = 0$ and consequently S is an independent set. ■

Theorem Let V be an inner product space and let $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ be any independent in V. Then one may construct orthogonal vectors $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_k$ in V such that for each $n = 1, 2, ..., k$ the set $\{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n\}$ is a basis for the subspace spanned by $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$.

Proof: The vectors $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n$ will be obtained by means of construction known the **Gram-Schmidt orthogonalisation process**. First let $\mathbf{y}_1 = \mathbf{x}_1$. The other vectors are then given inductively as follows: Suppose that $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_m$ $1 \leq m < n$ have been chosen so that for every i

$$\{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_i\}, \quad 1 \leq i \leq m$$

is an orthogonal basis for the subspace of V that is spanned by $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_i$. To construct the next vector $\mathbf{y}_{i+1}$, let

$$\mathbf{y}_{i+1} = \mathbf{x}_{i+1} - \sum_{j=1}^{i} \frac{\left(\mathbf{x}_{i+1} \bullet \mathbf{y}_j\right)}{\|\mathbf{y}_j\|^2} \mathbf{y}_j.$$

Then $\mathbf{y}_{i+1} \neq \mathbf{0}$, as otherwise $\mathbf{x}_{i+1}$ is a linear combination of $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_i$ and hence a linear combination of $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_i$. Furthermore, if $1 \leq l \leq i$ we have

$$(\mathbf{y}_{i+1} \bullet \mathbf{y}_l) = (\mathbf{x}_{i+1} \bullet \mathbf{y}_l) - \sum_{j=1}^{i} \frac{\left(\mathbf{x}_{i+1} \bullet \mathbf{y}_j\right)}{\|\mathbf{x}_j\|^2} (\mathbf{x}_j \bullet \mathbf{y}_l) = (\mathbf{x}_{i+1} \bullet \mathbf{y}_l) - \left(\mathbf{x}_{i+1} \bullet \mathbf{y}_j\right) = 0.$$

Hence $\{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_i, \mathbf{y}_{i+1}\}$ is an orthogonal set consisting of $i+1$ non-zero vectors in the subspace spanned by $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_i, \mathbf{x}_{i+1}$. By the previous theorem the set $\{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_i, \mathbf{y}_{i+1}\}$ is linearly independent and consequently is a basis for this subspace. Hence we can construct the vectors $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_k$ one after the other. For example, when $k = 4$ we have

$$\begin{aligned}
\mathbf{y}_1 &= \mathbf{x}_1 \\
\mathbf{y}_2 &= \mathbf{x}_2 - \frac{(\mathbf{x}_2 \bullet \mathbf{y}_1)}{\|\mathbf{y}_1\|^2} \mathbf{y}_1 \\
\mathbf{y}_3 &= \mathbf{x}_3 - \frac{(\mathbf{x}_3 \bullet \mathbf{y}_1)}{\|\mathbf{y}_1\|^2} \mathbf{y}_1 - \frac{(\mathbf{x}_3 \bullet \mathbf{y}_2)}{\|\mathbf{y}_2\|^2} \mathbf{y}_2 \\
\mathbf{y}_4 &= \mathbf{x}_4 - \frac{(\mathbf{x}_4 \bullet \mathbf{y}_1)}{\|\mathbf{y}_1\|^2} \mathbf{y}_1 - \frac{(\mathbf{x}_4 \bullet \mathbf{y}_2)}{\|\mathbf{y}_2\|^2} \mathbf{y}_2 - \frac{(\mathbf{x}_4 \bullet \mathbf{y}_3)}{\|\mathbf{y}_3\|^2} \mathbf{y}_3.
\end{aligned}$$

■

Corollary Every finite-dimensional inner product space has an orthonormal basis.

This is obvious as if $\{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n\}$ is a basis of V, a finite dimensional inner product space, applying the Gram-Schmidt process we can construct an orthogonal basis $\{\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_n\}$ and then obtain the orthonormal basis by simply replace each vector $\mathbf{y}_k$ by $\frac{1}{\|\mathbf{y}_k\|}\mathbf{y}_k$.

Exercises:

1. Let V be a vector space and $(\bullet)$ an inner product on V. a) Show that $(\mathbf{0} \bullet \mathbf{x}) = 0$ for all $\mathbf{x}$ in V. b) Show that if $(\mathbf{x} \bullet \mathbf{y}) = 0$ for all $\mathbf{y}$ in V, then $\mathbf{x} = \mathbf{0}$.

2. Let V be a vector space over F. Show that the sum of two inner products on V is an inner product on V. Is the difference of two inner products an inner product? Show that a positive multiple of an inner product is an inner product.

3. Let $(\bullet)$ be the standard inner product on $\mathbb{R}^2$. a) Let $\mathbf{x} = (1, 2)^{/}$ and $\mathbf{y} = (-1, 1)^{/}$. If $\mathbf{z}$ is a vector such that $(\mathbf{x} \bullet \mathbf{z}) = -1$ and $(\mathbf{y} \bullet \mathbf{z}) = 3$, find $\mathbf{z}$. b) Show that for any $\mathbf{x}$ in $\mathbb{R}^2$ we have that $\mathbf{x} = (\mathbf{x} \bullet \mathbf{e}_1)\, \mathbf{e}_1 + (\mathbf{x} \bullet \mathbf{e}_2)\, \mathbf{e}_2$.