1

Vector spaces

In this chapter we study vector spaces and their basic properties and structures. We start by stating the definition and a discussion of the examples of vector spaces. We next introduce the notions of subspaces, linear dependence, bases, coordinates, and dimensionality. We then consider dual spaces, direct sums, and quotient spaces. Finally we cover normed vector spaces.

1.1 Vector spaces

A *vector space* is a non-empty set consisting of elements called *vectors* which can be added and multiplied by some quantities called *scalars*. In this section, we start with a study of vector spaces.

1.1.1 Fields

The scalars to operate on vectors in a vector space are required to form a *field*, which may be denoted by $\mathbb{F}$, where two operations, usually called addition, denoted by '$+$', and multiplication, denoted by '$\cdot$' or omitted, over $\mathbb{F}$ are performed between scalars, such that the following axioms are satisfied.

(1) (Closure) If $a, b \in \mathbb{F}$, then $a + b \in \mathbb{F}$ and $ab \in \mathbb{F}$.
(2) (Commutativity) For $a, b \in \mathbb{F}$, there hold $a + b = b + a$ and $ab = ba$.
(3) (Associativity) For $a, b, c \in \mathbb{F}$, there hold $(a + b) + c = a + (b + c)$ and $a(bc) = (ab)c$.
(4) (Distributivity) For $a, b, c \in \mathbb{F}$, there hold $a(b + c) = ab + ac$.
(5) (Existence of zero) There is a scalar, called zero, denoted by 0, such that $a + 0 = a$ for any $a \in \mathbb{F}$.
(6) (Existence of unity) There is a scalar different from zero, called one, denoted by 1, such that $1a = a$ for any $a \in \mathbb{F}$.

(7) (Existence of additive inverse) For any $a \in \mathbb{F}$, there is a scalar, denoted by $-a$ or $(-a)$, such that $a + (-a) = 0$.

(8) (Existence of multiplicative inverse) For any $a \in \mathbb{F} \setminus \{0\}$, there is a scalar, denoted by a^{-1}, such that $aa^{-1} = 1$.

It is easily seen that zero, unity, additive and multiplicative inverses are all unique. Besides, a field consists of at least two elements.

With the usual addition and multiplication, the sets of rational numbers, real numbers, and complex numbers, denoted by $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, respectively, are all fields. These fields are infinite fields. However, the set of integers, $\mathbb{Z}$, is not a field because there is a lack of multiplicative inverses for its non-unit elements.

Let p be a prime ($p = 2, 3, 5, \dots$) and set $p\mathbb{Z} = \{n \in \mathbb{Z} \,|\, n = kp, k \in \mathbb{Z}\}$. Classify $\mathbb{Z}$ into the so-called cosets modulo $p\mathbb{Z}$, that is, some non-overlapping subsets of $\mathbb{Z}$ represented as $[i]$ ($i \in \mathbb{Z}$) such that

$$[i] = \{j \in \mathbb{Z} \,|\, i - j \in p\mathbb{Z}\}. \tag{1.1.1}$$

It is clear that $\mathbb{Z}$ is divided into exactly p cosets, $[0], [1], \dots, [p-1]$. Use $\mathbb{Z}_p$ to denote the set of these cosets and pass the additive and multiplicative operations in $\mathbb{Z}$ over naturally to the elements in $\mathbb{Z}_p$ so that

$$[i] + [j] = [i + j], \quad [i][j] = [ij]. \tag{1.1.2}$$

It can be verified that, with these operations, $\mathbb{Z}_p$ becomes a field with its obvious zero and unit elements, $[0]$ and $[1]$. Of course, $p[1] = [1] + \cdots + [1]$ (p terms)$= [p] = [0]$. In fact, p is the smallest positive integer whose multiplication with unit element results in zero element. A number of such a property is called the *characteristic of the field*. Thus, $\mathbb{Z}_p$ is a *field of characteristic* p. For $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, since no such integer exists, we say that these fields are of *characteristic* 0.

1.1.2 Vector spaces

Let $\mathbb{F}$ be a field. Consider the set of n-tuples, denoted by $\mathbb{F}^n$, with elements called vectors arranged in row or column forms such as

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \quad \text{or} \quad (a_1, \dots, a_n) \quad \text{where } a_1, \dots, a_n \in \mathbb{F}. \tag{1.1.3}$$

Furthermore, we can define the addition of two vectors and the scalar multiplication of a vector by a scalar following the rules such as

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}, \tag{1.1.4}$$

$$\alpha \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \alpha a_1 \\ \vdots \\ \alpha a_n \end{pmatrix} \text{ where } \alpha \in \mathbb{F}. \tag{1.1.5}$$

The set $\mathbb{F}^n$, modeled over the field $\mathbb{F}$ and equipped with the above operations, is a prototype example of a vector space.

More generally, we say that a set U is a *vector space over a field* $\mathbb{F}$ if U is non-empty and there is an operation called *addition*, denoted by '+', between the elements of U, called *vectors*, and another operation called *scalar multiplication* between elements in $\mathbb{F}$, called *scalars*, and vectors, such that the following axioms hold.

(1) (Closure) For $u, v \in U$, we have $u + v \in U$. For $u \in U$ and $a \in \mathbb{F}$, we have $au \in U$.
(2) (Commutativity) For $u, v \in U$, we have $u + v = v + u$.
(3) (Associativity of addition) For $u, v, w \in U$, we have $u + (v + w) = (u + v) + w$.
(4) (Existence of zero vector) There is a vector, called zero and denoted by 0, such that $u + 0 = u$ for any $u \in U$.
(5) (Existence of additive inverse) For any $u \in U$, there is a vector, denoted as $(-u)$, such that $u + (-u) = 0$.
(6) (Associativity of scalar multiplication) For any $a, b \in \mathbb{F}$ and $u \in U$, we have $a(bu) = (ab)u$.
(7) (Property of unit scalar) For any $u \in U$, we have $1u = u$.
(8) (Distributivity) For any $a, b \in \mathbb{F}$ and $u, v \in U$, we have $(a+b)u = au+bu$ and $a(u + v) = au + av$.

As in the case of the definition of a field, we see that it readily follows from the definition that zero vector and additive inverse vectors are all unique in a vector space. Besides, any vector multiplied by zero scalar results in zero vector. That is, $0u = 0$ for any $u \in U$.

Other examples of vector spaces (with obviously defined vector addition and scalar multiplication) include the following.

(1) The set of all polynomials with coefficients in $\mathbb{F}$ defined by

$$\mathcal{P} = \{a_0 + a_1 t + \cdots + a_n t^n \,|\, a_0, a_1, \ldots, a_n \in \mathbb{F}, n \in \mathbb{N}\}, \tag{1.1.6}$$

where t is a variable parameter.

(2) The set of all real-valued continuous functions over the interval $[a, b]$ for $a, b \in \mathbb{R}$ and $a < b$ usually denoted by $C[a, b]$.

(3) The set of real-valued solutions to the differential equation

$$a_n \frac{\mathrm{d}^n x}{\mathrm{d}t^n} + \cdots + a_1 \frac{\mathrm{d}x}{\mathrm{d}t} + a_0 x = 0, \quad a_0, a_1, \ldots, a_n \in \mathbb{R}. \tag{1.1.7}$$

(4) In addition, we can also consider the set of arrays of scalars in $\mathbb{F}$ consisting of m rows of vectors in $\mathbb{F}^n$ or n columns of vectors in $\mathbb{F}^m$ of the form

$$(a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \tag{1.1.8}$$

where $a_{ij} \in \mathbb{F}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, called an m by n or $m \times n$ *matrix* and each a_{ij} is called an *entry* or *component* of the matrix. The set of all $m \times n$ matrices with entries in $\mathbb{F}$ may be denoted by $\mathbb{F}(m, n)$. In particular, $\mathbb{F}(m, 1)$ or $\mathbb{F}(1, n)$ is simply $\mathbb{F}^m$ or $\mathbb{F}^n$. Elements in $\mathbb{F}(n, n)$ are also called *square matrices*.

1.1.3 Matrices

Here we consider some of the simplest manipulations on, and properties of, matrices.

Let A be the matrix given in (1.1.8). Then A^t, called the *transpose* of A, is defined to be

$$A^t = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix}. \tag{1.1.9}$$

Of course, $A^t \in \mathbb{F}(n, m)$. Simply put, A^t is a matrix obtained from taking the row (column) vectors of A to be its corresponding column (row) vectors.

For $A \in \mathbb{F}(n, n)$, we say that A is *symmetric* if $A = A^t$, or *skew-symmetric* or *anti-symmetric* if $A^t = -A$. The sets of symmetric and anti-symmetric matrices are denoted by $\mathbb{F}_S(n, n)$ and $\mathbb{F}_A(n, n)$, respectively.

It is clear that $(A^t)^t = A$.

It will now be useful to introduce the notion of *dot product*. For any two vectors $u = (a_1, \dots, a_n)$ and $v = (b_1, \dots, b_n)$ in $\mathbb{F}^n$, their dot product $u \cdot v \in \mathbb{F}$ is defined to be

$$u \cdot v = a_1 b_1 + \cdots + a_n b_n. \tag{1.1.10}$$

The following properties of dot product can be directly examined.

(1) (Commutativity) $u \cdot v = v \cdot u$ for any $u, v \in \mathbb{F}^n$.
(2) (Associativity and homogeneity) $u \cdot (av + bw) = a(u \cdot v) + b(u \cdot w)$ for any $u, v, w \in \mathbb{F}^n$ and $a, b \in \mathbb{F}$.

With the notion of dot product, we can define the *product of two matrices* $A \in \mathbb{F}(m, k)$ and $B \in \mathbb{F}(k, n)$ by

$$C = (c_{ij}) = AB, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \tag{1.1.11}$$

where c_{ij} is the dot product of the ith row of A and the jth column of B. Thus $AB \in \mathbb{F}(m, n)$.

Alternatively, if we use u, v to denote column vectors in $\mathbb{F}^n$, then

$$u \cdot v = u^t v. \tag{1.1.12}$$

That is, the dot product of u and v may be viewed as a matrix product of the $1 \times n$ matrix u^t and $n \times 1$ matrix v as well.

Matrix product (or *matrix multiplication*) enjoys the following properties.

(1) (Associativity of scalar multiplication) $a(AB) = (aA)B = A(aB)$ for any $a \in \mathbb{F}$ and any $A \in \mathbb{F}(m, k)$, $B \in \mathbb{F}(k, n)$.
(2) (Distributivity) $A(B + C) = AB + AC$ for any $A \in \mathbb{F}(m, k)$ and $B, C \in \mathbb{F}(k, n)$; $(A + B)C = AC + BC$ for any $A, B \in \mathbb{F}(m, k)$ and $C \in \mathbb{F}(k, n)$.
(3) (Associativity) $A(BC) = (AB)C$ for any $A \in \mathbb{F}(m, k)$, $B \in \mathbb{F}(k, l)$, $C \in \mathbb{F}(l, n)$.

Alternatively, if we express $A \in \mathbb{F}(m, k)$ and $B \in \mathbb{F}(k, n)$ as made of m row vectors and n column vectors, respectively, rewritten as

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix}, \quad B = (B_1, \dots, B_n), \tag{1.1.13}$$

then, formally, we have

$$AB = \begin{pmatrix} A_1 \cdot B_1 & A_1 \cdot B_2 & \cdots & A_1 \cdot B_n \\ A_2 \cdot B_1 & A_2 \cdot B_2 & \cdots & A_2 \cdot B_n \\ \cdots & \cdots & \cdots & \cdots \\ A_m \cdot B_1 & A_m \cdot B_2 & \cdots & A_m \cdot B_n \end{pmatrix}$$
$$= \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} (B_1, \ldots, B_n)$$
$$= \begin{pmatrix} A_1 B_1 & A_1 B_2 & \cdots & A_1 B_n \\ A_2 B_1 & A_2 B_2 & \cdots & A_2 B_n \\ \cdots & \cdots & \cdots & \cdots \\ A_m B_1 & A_m B_2 & \cdots & A_m B_n \end{pmatrix}, \tag{1.1.14}$$

which suggests that matrix multiplication may be carried out with legitimate multiplications executed over appropriate matrix blocks.

If $A \in \mathbb{F}(m, k)$ and $B \in \mathbb{F}(k, n)$, then $A^t \in \mathbb{F}(k, m)$ and $B^t \in \mathbb{F}(n, k)$ so that $B^t A^t \in \mathbb{F}(n, m)$. Regarding how AB and $B^t A^t$ are related, here is the conclusion.

Theorem 1.1 *For $A \in \mathbb{F}(m, k)$ and $B \in \mathbb{F}(k, n)$, there holds*

$$(AB)^t = B^t A^t. \tag{1.1.15}$$

The proof of this basic fact is assigned as an exercise.

Other matrices in $\mathbb{F}(n, n)$ having interesting properties include the following.

(1) *Diagonal matrices* are of the form $A = (a_{ij})$ with $a_{ij} = 0$ whenever $i \neq j$. The set of diagonal matrices is denoted as $\mathbb{F}_D(n, n)$.
(2) *Lower triangular matrices* are of the form $A = (a_{ij})$ with $a_{ij} = 0$ whenever $j > i$. The set of lower triangular matrices is denoted as $\mathbb{F}_L(n, n)$.
(3) *Upper triangular matrices* are of the form $A = (a_{ij})$ with $a_{ij} = 0$ whenever $i > j$. The set of upper triangular matrices is denoted as $\mathbb{F}_U(n, n)$.

There is a special element in $\mathbb{F}(n, n)$, called the *identity matrix*, or *unit matrix*, and denoted by I_n, or simply I, which is a diagonal matrix whose diagonal entries are all 1 (unit scalar) and off-diagonal entries are all 0. For any $A \in \mathbb{F}(n, n)$, we have $AI = IA = A$.

Definition 1.2 A matrix $A \in \mathbb{F}(n, n)$ is called *invertible* or *nonsingular* if there is some $B \in \mathbb{F}(n, n)$ such that

$$AB = BA = I. \tag{1.1.16}$$

In this situation, B is unique (cf. Exercise 1.1.7) and called the *inverse* of A and denoted by A^{-1}.

If $A, B \in \mathbb{F}(n, n)$ are such that $AB = I$, then we say that A is a *left inverse* of B and B a *right inverse* of A. It can be shown that a left or right inverse is simply the inverse. In other words, if A is a left inverse of B, then both A and B are invertible and the inverses of each other.

If $A \in \mathbb{R}(n, n)$ enjoys the property $AA^t = A^tA = I$, then A is called an *orthogonal matrix*. For $A = (a_{ij}) \in \mathbb{C}(m, n)$, we adopt the notation $\overline{A} = (\overline{a}_{ij})$ for taking the complex conjugate of A and use $A^\dagger$ to denote taking the complex conjugate of the transpose of A, $A^\dagger = \overline{A}^t$, which is also commonly referred to as taking the *Hermitian conjugate* of A. If $A \in \mathbb{C}(n, n)$, we say that A is *Hermitian symmetric*, or simply *Hermitian*, if $A^\dagger = A$, and *skew-Hermitian* or *anti-Hermitian*, if $A^\dagger = -A$. If $A \in \mathbb{C}(n, n)$ enjoys the property $AA^\dagger = A^\dagger A = I$, then A is called a *unitary matrix*. We will see the importance of these notions later.

Exercises

1.1.1 Show that it follows from the definition of a field that zero, unit, additive, and multiplicative inverse scalars are all unique.

1.1.2 Let $p \in \mathbb{N}$ be a prime and $[n] \in \mathbb{Z}_p$. Find $-[n]$ and prove the existence of $[n]^{-1}$ when $[n] \neq [0]$. In $\mathbb{Z}_5$, find $-[4]$ and $[4]^{-1}$.

1.1.3 Show that it follows from the definition of a vector space that both zero and additive inverse vectors are unique.

1.1.4 Prove the associativity of matrix multiplication by showing that $A(BC) = (AB)C$ for any $A \in \mathbb{F}(m, k)$, $B \in \mathbb{F}(k, l)$, $C \in \mathbb{F}(l, n)$.

1.1.5 Prove Theorem 1.1.

1.1.6 Let $A \in \mathbb{F}(n, n)$ $(n \geq 2)$ and rewrite A as

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}, \tag{1.1.17}$$

where $A_1 \in \mathbb{F}(k, k)$, $A_2 \in \mathbb{F}(k, l)$, $A_3 \in \mathbb{F}(l, k)$, $A_4 \in \mathbb{F}(l, l)$, $k, l \geq 1$, $k + l = n$. Show that

$$A^t = \begin{pmatrix} A_1^t & A_3^t \\ A_2^t & A_4^t \end{pmatrix}. \tag{1.1.18}$$

1.1.7 Prove that the inverse of an invertible matrix is unique by showing the fact that if $A, B, C \in \mathbb{F}(n, n)$ satisfy $AB = I$ and $CA = I$ then $B = C$.

1.1.8 Let $A \in \mathbb{C}(n, n)$. Show that A is Hermitian if and only if $\mathrm{i}A$ is anti-Hermitian.

1.2 Subspaces, span, and linear dependence

Let U be a vector space over a field $\mathbb{F}$ and $V \subset U$ a non-empty subset of U. We say that V is a *subspace* of U if V is a vector space over $\mathbb{F}$ with the inherited addition and scalar multiplication from U. It is worth noting that, when checking whether a subset V of a vector space U becomes a subspace, one only needs to verify the closure axiom (1) in the definition of a vector space since the rest of the axioms follow automatically as a consequence of (1).

The two trivial subspaces of U are those consisting only of zero vector, $\{0\}$, and U itself. A nontrivial subspace is also called a *proper subspace*.

Consider the subset $\mathcal{P}_n$ ($n \in \mathbb{N}$) of $\mathcal{P}$ defined by

$$\mathcal{P}_n = \{a_0 + a_1 t + \cdots + a_n t^n \,|\, a_0, a_1, \ldots, a_n \in \mathbb{F}\}. \tag{1.2.1}$$

It is clear that $\mathcal{P}_n$ is a subspace of $\mathcal{P}$ and $\mathcal{P}_m$ is subspace of $\mathcal{P}_n$ when $m \leq n$.

Consider the set S_a of all vectors $(x_1, \ldots, x_n)$ in $\mathbb{F}^n$ satisfying the equation

$$x_1 + \cdots + x_n = a, \tag{1.2.2}$$

where $a \in \mathbb{F}$. Then S_a is a subspace of $\mathbb{F}^n$ if and only if $a = 0$.

Let $u_1, \ldots, u_k$ be vectors in U. The *linear span* of $\{u_1, \ldots, u_k\}$, denoted by $\mathrm{Span}\{u_1, \ldots, u_k\}$, is the subspace of U defined by

$$\mathrm{Span}\{u_1, \ldots, u_k\} = \{u \in U \,|\, u = a_1 u_1 + \cdots + a_k u_k,\ a_1, \ldots, a_k \in \mathbb{F}\}. \tag{1.2.3}$$

Thus, if $u \in \mathrm{Span}\{u_1, \ldots, u_k\}$, then there are $a_1, \ldots, a_k \in \mathbb{F}$ such that

$$u = a_1 u_1 + \cdots + a_k u_k. \tag{1.2.4}$$

We also say that u is *linearly spanned* by $u_1, \ldots, u_k$ or *linearly dependent* on $u_1, \ldots, u_k$. Therefore, zero vector 0 is linearly dependent on any finite set of vectors.

If $U = \mathrm{Span}\{u_1, \ldots, u_k\}$, we also say that U is *generated* by the vectors $u_1, \ldots, u_k$.

For $\mathcal{P}_n$ defined in (1.2.1), we have $\mathcal{P}_n = \mathrm{Span}\{1, t, \ldots, t^n\}$. Thus $\mathcal{P}_n$ is generated by $1, t, \ldots, t^n$. Naturally, for two elements p, q in

$\mathcal{P}_n$, say $p(t) = a_0 + a_1 t + \cdots + a_n t^n$, $q(t) = b_0 + b_1 t + \cdots + b_n t^n$, we identify p and q if and only if all the coefficients of p and q of like powers of t coincide in $\mathbb{F}$, or, $a_i = b_i$ for all $i = 0, 1, \ldots, n$.

In $\mathbb{F}^n$, define

$$e_1 = (1, 0, 0, \ldots, 0), \quad e_2 = (0, 1, 0, \ldots, 0), \quad e_n = (0, 0, \ldots, 0, 1). \tag{1.2.5}$$

Then $\mathbb{F}^n = \text{Span}\{e_1, e_2, \ldots, e_n\}$ and $\mathbb{F}^n$ is generated by $e_1, e_2, \ldots, e_n$.

Thus, for S_0 defined in (1.2.2), we have

$$\begin{aligned}(x_1, x_2, \ldots, x_n) &= -(x_2 + \cdots + x_n)e_1 + x_2 e_2 + \cdots + x_n e_n \\ &= x_2(e_2 - e_1) + \cdots + x_n(e_n - e_1),\end{aligned} \tag{1.2.6}$$

where $x_2, \ldots, x_n$ are arbitrarily taken from $\mathbb{F}$. Therefore

$$S_0 = \text{Span}\{e_2 - e_1, \ldots, e_n - e_1\}. \tag{1.2.7}$$

For $\mathbb{F}(m, n)$, we define $M_{ij} \in \mathbb{F}(m, n)$ to be the vector such that all its entries vanish except that its entry at the position (i, j) (at the ith row and jth column) is 1, $i = 1, \ldots, m$, $j = 1, \ldots, n$. We have

$$\mathbb{F}(m, n) = \text{Span}\{M_{ij} \mid i = 1, \ldots, m, j = 1, \ldots, n\}. \tag{1.2.8}$$

The notion of spans can be extended to cover some useful situations. Let U be a vector space and S be a (finite or infinite) subset of U. Define

$$\begin{aligned}\text{Span}(S) = &\text{ the set of linear combinations} \\ &\text{of all possible finite subsets of } S.\end{aligned} \tag{1.2.9}$$

It is obvious that $\text{Span}(S)$ is a subspace of U. If $U = \text{Span}(S)$, we say that U is spanned or generated by the set of vectors S.

As an example, we have

$$\mathcal{P} = \text{Span}\{1, t, \ldots, t^n, \ldots\}. \tag{1.2.10}$$

Alternatively, we can also express $\mathcal{P}$ as

$$\mathcal{P} = \cup_{n=0}^{\infty} \mathcal{P}_n. \tag{1.2.11}$$

The above discussion motivates the following formal definition.

Definition 1.3 Let $u_1, \ldots, u_m$ be m vectors in the vector space U over a field $\mathbb{F}$. We say that these vectors are *linearly dependent* if one of them may be written as a linear span of the rest of them or linearly dependent on the rest

of them. Or equivalently, $u_1, \dots, u_m$ are linearly dependent if there are scalars $a_1, \dots, a_m \in \mathbb{F}$ where $(a_1, \dots, a_m) \neq (0, \dots, 0)$ such that

$$a_1 u_1 + \cdots + a_m u_m = 0. \tag{1.2.12}$$

Otherwise $u_1, \dots, u_m$ are called *linearly independent*. In this latter case, the only possible vector $(a_1, \dots, a_m) \in \mathbb{F}^n$ to make (1.2.12) fulfilled is the zero vector, $(0, \dots, 0)$.

To proceed further, we need to consider the following system of linear equations

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots \quad \cdots \quad \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0, \end{cases} \tag{1.2.13}$$

over $\mathbb{F}$ with unknowns $x_1, \dots, x_n$.

Theorem 1.4 *In the system (1.2.13), if $m < n$, then the system has a nontrivial solution $(x_1, \dots, x_n) \neq (0, \dots, 0)$.*

Proof We prove the theorem by using induction on $m + n$.

The beginning situation is $m + n = 3$ when $m = 1$ and $n = 2$. It is clear that we always have a nontrivial solution.

Assume that the statement of the theorem is true when $m + n \leq k$ where $k \geq 3$.

Let $m + n = k + 1$. If $k = 3$, the condition $m < n$ implies $m = 1, n = 3$ and the existence of a nontrivial solution is obvious. Assume then $k \geq 4$. If all the coefficients of the variable x_1 in (1.2.13) are zero, i.e. $a_{11} = \cdots = a_{m1} = 0$, then $x_1 = 1, x_2 = \cdots = x_n = 0$ is a nontrivial solution. So we may assume one of the coefficients of x_1 is nonzero. Without loss of generality, we assume $a_{11} \neq 0$. If $m = 1$, there is again nothing to show. Assume $m \geq 2$. Dividing the first equation in (1.2.13) by a_{11} if necessary, we can further assume $a_{11} = 1$. Then, adding the $(-a_{i1})$-multiple of the first equation into the ith equation, in (1.2.13), for $i = 2, \dots, m$, we arrive at

$$\begin{cases} x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ \quad\; b_{22}x_2 + \cdots + b_{2n}x_n = 0, \\ \quad\; \cdots\cdots\cdots\cdots\cdots \quad \cdots \quad \cdots \\ \quad\; b_{m2}x_2 + \cdots + b_{mn}x_n = 0. \end{cases} \tag{1.2.14}$$

The system below the first equation in (1.2.14) contains $m - 1$ equations and $n - 1$ unknowns $x_2, \dots, x_n$. Of course, $m - 1 < n - 1$. So, in view of the