CHAPTER 2

SYSTEMS OF LINEAR EQUATIONS ◊

2.1 Matrices

2.1.1 Vector spaces

Definition 2.1.1 Let X be a nonempty set and K be a field. We say that X is a *vector* (or *linear*) *space over the field* K if :

- X is provided with a binary operation, called *addition*, that maps any $(\boldsymbol{x},\boldsymbol{y}) \in X \times X$ into one and only one element of X, denoted $\boldsymbol{x}+\boldsymbol{y}$, satisfying the following statements

(2.1.1) (commutative property) $\forall\,(\boldsymbol{x},\boldsymbol{y}) \in X \times X \qquad \boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$

(2.1.2) (associative property) $\forall(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \in X^3 \qquad \boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z}) = (\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}$

◊ A. Maceri, *Systems of linear equations*, e-ISBN 978-88-85929-60-9, © Accademica 2019

(2.1.3) there exists an element $\mathbf{0} \in X$, called the *zero element*, such that $\forall\, \boldsymbol{x} \in X \qquad \mathbf{0} + \boldsymbol{x} = \boldsymbol{x}$

(2.1.4) $\forall\, \boldsymbol{x} \in X$ there exists an element $-\boldsymbol{x} \in X$, called the *opposite element* of $\boldsymbol{x}$, with the property that $\boldsymbol{x} + (-\boldsymbol{x}) = \mathbf{0}$;

- X is provided with an operation $\cdot$, called *scalar multiplication*, that maps any $(\alpha, \boldsymbol{x}) \in K \times X$ into one and only one element of X , called product of α and $\boldsymbol{x}$, denoted $\alpha \cdot \boldsymbol{x}$ (or $\alpha\boldsymbol{x}$), satisfying the following statements

(2.1.5) $\forall\, (\alpha, \beta, \boldsymbol{x}) \in K \times K \times X \qquad \alpha \cdot (\beta \cdot \boldsymbol{x}) = (\alpha \cdot \beta) \cdot \boldsymbol{x}$

(2.1.6) X contains an element $\mathbf{1} \neq \mathbf{0}$, called *identity element*, such that $\forall\, \boldsymbol{x} \in X \qquad \mathbf{1} \cdot \boldsymbol{x} = \boldsymbol{x}$;

- the operations of addition and multiplication obey the *distributive laws*

(2.1.7) $\forall\, (\alpha, \beta, \boldsymbol{x}) \in K \times K \times X \qquad (\alpha + \beta) \cdot \boldsymbol{x} = \alpha \cdot \boldsymbol{x} + \beta \cdot \boldsymbol{x}$

(2.1.8) $\forall\, (\alpha, \boldsymbol{x}, \boldsymbol{y}) \in K \times X \times X \qquad \alpha \cdot (\boldsymbol{x} + \boldsymbol{y}) = \alpha \cdot \boldsymbol{x} + \alpha \cdot \boldsymbol{y}$.

The elements of X are called *vectors* (or *points*), while the elements of K are called *scalars*. If $K = \mathbb{R}$, X is called *real vector* (or *linear*) *space.* ⋄

A very important example of vector space over a field is the *vector space* $\mathbb{R}^n$ *over the real field*. Precisely, for each positive integer n , let

$\mathbb{R}^n$ be the set of all ordered n-tuples

$$\boldsymbol{x} = (x_1, x_2, \dots, x_n),$$

where $x_1, x_2, \dots, x_n$ are real numbers. For each $\boldsymbol{y} = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ and for each $\alpha \in \mathbb{R}$, we put

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, \dots, x_n + y_n)$$

$$\alpha\boldsymbol{x} = (\alpha x_1, \dots, \alpha\, x_n)$$

so that $\boldsymbol{x} + \boldsymbol{y} \in \mathbb{R}^n$ and $\alpha\boldsymbol{x} \in \mathbb{R}^n$. We easily verify that such operations (addition of vectors and multiplication of a vector by a scalar) satisfy the commutative, associative and distributive laws, as well as that the zero element of $\mathbb{R}^n$ (sometimes called the *origin* or the *null vector*) is the point $\mathbf{0}$, all of whose coordinates are 0. Even the existence of the opposite element and the identity element is easily verified.

Definition 2.1.2 Let

$$x_1 = (x_{1,1}, x_{1,2}, \dots, x_{1,n})$$

$$x_2 = (x_{2,1}, x_{2,2}, \dots, x_{2,n})$$

$$\dots$$

$$x_k = (x_{k,1}, x_{k,2}, \dots, x_{k,n})$$

Be $k \in \mathbb{N}$ vectors of $\mathbb{R}^n$, and $c_1, c_2, \dots, c_k$ be k real numbers. The vector of $\mathbb{R}^n$

$$x = c_1 x_1 + c_2 x_2 + \cdots + c_k x_k \tag{2.1.9}$$

is called the *linear combination* of the vectors $\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_k}$ according the *coefficients* $c_1, c_2, \dots, c_k$. ◊

Remark 2.1.1 Obviously, with the usual notation $x = (x_1, x_2, \dots, x_n)$, the vector equality (2.1.9) is equivalent to system of k scalar equalities

(2.1.10)

$$x_1 = c_1 x_{1,1} + c_2 x_{2,1} + \cdots + c_k x_{k,1}$$

$$x_2 = c_1 x_{1,2} + c_2 x_{2,2} + \cdots + c_k x_{k,2}$$

$$\dots$$

$$x_n = c_1 x_{1,n} + c_2 x_{2,n} + \cdots + c_k x_{k,n} .$$ ◊

Remark 2.1.2 Obviously, the linear combination of any vectors $\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_k}$ according the coefficients $0,0,\dots,0$ is the null vector. ◊

Definition 2.1.3 Let $\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_k}$ be $k \in \mathbb{N}$ vectors of $\mathbb{R}^n$. If

$$c_1 x_1 + c_2 x_2 + \cdots + c_k x_k = 0 \tag{2.1.11}$$

only happens when $c_1 = c_2 = \cdots = c_k = 0$, then the vectors $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_k}$ are called *linearly independent*. ◊

Definition 2.1.4 Let $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_k}$ be $k \in \mathbb{N}$ vectors of $\mathbb{R}^n$. If there exists a $(c_1, c_2, \ldots, c_k) \in \mathbb{R}^k - \{\mathbf{0}\}$ such that

$$c_1x_1 + c_2x_2 + \cdots + c_kx_k = 0 \ ,$$

then the vectors $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_k}$ are called *linearly dependent*. ◊

Remark 2.1.3 Clearly, the null vector is linearly dependent and every nonnull vector is linearly independent. ◊

Theorem 2.1.1 *Let X be a real vector space and $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_k}$ $(k \in \mathbb{N})$ be a system of vectors of X. The following statements are equivalent:*

1) *the vectors are linearly dependent*
2) *one vector is a linear combination of the others*.

Proof. 1) ⇒ 2). Let $c_1 \neq 0$ in (2.1.11). Then

$$x_1 = -\frac{c_2}{c_1}\, x_2 - \cdots - \frac{c_k}{c_1}\, x_k \ .$$

2) $\Rightarrow$ 1). Suppose that one of the vectors $\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_k}$ $(k \in \mathbb{N})$ is a linear combination of the others. Then the vectors $\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_k}$ are linearly dependent. ◊

Definition 2.1.5 The real vector space X is called finite-dimensional and the number n is called the dimension of the space if there exist n linearly independent vectors in X, while any $n+1$ vectors in X are linearly dependent. If the space contains linearly independent systems of an arbitrary number of vectors, then it is called infinite-dimensional. ◊

Definition 2.1.6 Let X be a real vector space. We call *basis* of X a system of vectors of X if

- such vectors are linearly independent
- the vector space X consists of all their linear combinations. ◊

Theorem 2.1.2 *The coordinate vectors*

(2.1.12)
$$\boldsymbol{e_1} = (1,0,\dots,0)$$
$$\boldsymbol{e_2} = (0,1,\dots,0)$$
$$\dots$$
$$\boldsymbol{e_n} = (0,0,\dots,1)$$

are a basis of the real vector space $\mathbb{R}^n$.

Proof. Since the vector equation $c_1e_1 + c_2e_2 + \cdots + c_ne_n = 0$ implies the scalar equations

$$c_1 = 0$$
$$c_2 = 0$$
$$\ldots$$
$$c_n = 0\ ,$$

the coordinate vectors $e_1, e_2, \ldots, e_n$ are linearly independent.
Let $x = (x_1, x_2, \ldots, x_n)$ be any vector of $\mathbb{R}^n$. Since

$$x = x_1e_1 + x_2e_2 + \cdots + x_ne_n\ ,$$

the thesis is true. ◊

Remark 2.1.4 There is one and only one way to write a vector as a linear combination of the basis vectors. In fact, if

$$x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$$
$$x = c_1e_1 + c_2e_2 + \cdots + c_ne_n\ ,$$

then subtraction gives

$$0 = (x_1 - c_1)e_1 + (x_2 - c_2)e_2 + \cdots + (x_n - c_n)e_n\ .$$

By independence, every coefficient must be zero, hence $x_1 = c_1$, $x_2 = c_2$, … , $x_n = c_n$. ◊

Remark 2.1.5 If $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n$, then $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$, and the numbers x_1 , x_2 , …, x_n are called the *coordinates of* x *in the basis* $e_1, e_2, \dots, e_n$. ◊

2.1.2 Basics of matrices

Definition 2.1.7 Let $m, n \in \mathbb{N}$. A rectangular array of real numbers

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

is called a *real matrix*. When $m = n$, the matrix is called *square* and the number m , equal to n , is called its *order*. In the general case the matrix is called *rectangular* (of *dimension* $m \times n$). The numbers that constitute the matrix are called the *elements*. In the double-subscript notation for the elements, the first subscript always denotes the row and the second subscript the column containing the given element. ◊

Definition 2.1.8 A rectangular matrix consisting of a single column

$$\begin{bmatrix} a_{1j} \\ \cdots \\ a_{mj} \end{bmatrix}$$

is called *column matrix* or *column vector*. ◊

Definition 2.1.9 A rectangular matrix consisting of a single row

$$[a_{i1} \quad \cdots \quad a_{in}]$$

is called *row matrix* or *row vector*. ◊

Definition 2.1.10 A square matrix of order $n \in \mathbb{N}$, in which all the elements outside the main diagonal are zero

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix},$$

is called *diagonal matrix*. ◊

Definition 2.1.11 A square matrix of order $n \in \mathbb{N}$, in which the main diagonal consists entirely of units and all the others elements are zero

$$U = \begin{bmatrix} 1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 1 \end{bmatrix},$$

is called *unit matrix*. ◊

Definition 2.1.12 A square matrix in which all the elements below the main diagonal are zero

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

is called *upper triangular*. ◊

Definition 2.1.13 A square matrix in which all the elements above the main diagonal are zero

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

is called *lower triangular*. ◊

Definition 2.1.14 Let us consider the rectangular matrix of dimension $m \times n$

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} .$$

We call *transpose* of A the matrix of dimension $n \times m$

$$A^T = \begin{bmatrix} a_{11}^T & \dots & a_{1m}^T \\ \dots & \dots & \dots \\ a_{n1}^T & \dots & a_{nm}^T \end{bmatrix} ,$$

where $a_{ki}^T = a_{ik}$ $\forall i \in \{1,2,\dots,m\}$ and $\forall\, k \in \{1,2,\dots,n\}$. ◊

Definition 2.1.15 If a square matrix coincides with its transpose, then it is called *symmetric*. ◊

Remark 2.1.6 Let A any matrix of dimension $m \times n$. The transpose matrix A^T has as first row the first column of A ,…, as n-th row the n-th column of A . ◊

Definition 2.1.16 If a square matrix differs from its transpose by the factor -1 , then it is called *skew-symmetric*. ◊

Remark 2.1.7 In any symmetric matrix elements that are symmetrically placed with respect to the main diagonal are equal. In a skew-symmetric matrix elements any two elements that are that are symmetrically placed with respect to the main diagonal differ each other by the factor -1 and the diagonal elements are zero. ◊

2.1.3 Operations on matrices

Definition 2.1.17 Let

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots \\ b_{m1} & \dots & b_{mn} \end{bmatrix}$$

be two rectangular matrices, both of dimension $m \times n$. We call *sum* (or *addition*) of A and B, and denote $A + B$, the matrix, of the same dimension, whose elements are the sums of the corresponding elements of the given matrices:

$$A + B = \begin{bmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \dots & \dots & \dots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{bmatrix}. \ \diamond$$

Remark 2.1.8 According to definition 2.1.17, only rectangular matrices of equal dimension can be added. ◊

Remark 2.1.9 From the definition 2.1.17 it follows immediately that the matrix addition has the properties of commutativity $(A + B = B + A)$ and of associativity $((A + B) + C = A + (B + C))$. ◊

Definition 2.1.18 Let

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

be a rectangular matrix, of dimension $m \times n$. We call *product* (or *multiplication*) of A by $\alpha \in \mathbb{R}$, and denote αA , the matrix, of the same dimension, whose elements are obtained from the corresponding elements of A by multiplication by α :

$$\alpha A = \begin{bmatrix} \alpha a_{11} & \dots & \alpha a_{1n} \\ \dots & \dots & \dots \\ \alpha a_{m1} & \dots & \alpha a_{mn} \end{bmatrix} . \diamond$$

Remark 2.1.10 Let A , B be two rectangular matrices of equal dimension. It is easy to see that $\forall \alpha , \beta \in \mathbb{R}$

$$\alpha(A + B) = \alpha A + \alpha B \tag{2.1.13}$$

(2.1.14) $$(\alpha+\beta)A=\alpha\,A+\beta\,A$$

(2.1.15) $$(\alpha\,\beta)A=\alpha(\beta\,A)\ .\ \diamond$$

Definition 2.1.19 Let

$$A=\begin{bmatrix}a_{11} & \dots & a_{1n}\\ \dots & \dots & \dots\\ a_{m1} & \dots & a_{mn}\end{bmatrix},\ B=\begin{bmatrix}b_{11} & \dots & b_{1n}\\ \dots & \dots & \dots\\ b_{m1} & \dots & b_{mn}\end{bmatrix}$$

be two rectangular matrices, both of dimension $m\times n$. We call *difference* of A and B , and denote $A-B$, the matrix, of the same dimension

$$A-B=A+(-1)B\ .\ \diamond$$

Definition 2.1.20 Let

$$A=\begin{bmatrix}a_{11} & \dots & a_{1n}\\ \dots & \dots & \dots\\ a_{m1} & \dots & a_{mn}\end{bmatrix}$$

be a rectangular matrix of dimension $m\times n$,

$$B=\begin{bmatrix}b_{11} & \dots & b_{1q}\\ \dots & \dots & \dots\\ b_{n1} & \dots & b_{nq}\end{bmatrix}$$

be a rectangular matrix of dimension $n \times q$. We call *product* (or *multiplication*) of A and B , and denote $A\,B$, the matrix, of dimension $m \times q$

$$A\,B = \begin{bmatrix} c_{11} & \cdots & c_{1q} \\ \cdots & \cdots & \cdots \\ c_{m1} & \cdots & c_{mq} \end{bmatrix}$$

in which the element c_{ij} at the intersection of the i-th row and the j-th column is the sum of the products of the corresponding elements (of the i-th row and the j-th column):

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\,b_{kj} \qquad \forall\, i \in \{1, \dots, m\} \quad \text{and} \quad \forall\, j \in \{1, \dots, q\}\ .\ \diamond$$

Remark 2.1.11 According to definition 2.1.20, the operation of multiplication of two rectangular matrices can only be carried out when the number of columns of the first factor is equal to the number of rows of the second. ◊

Remark 2.1.12 It is easy to see that the matrix multiplication has the following properties

(2.1.16) $(A\,B)C = A(B\,C)$

(2.1.17) $(A + B)C = A\,C + B\,C$

(2.1.18) $A(B + C) = A\,B + A\,C$. ◊

Theorem 2.1.3 *If a rectangular matrix* A *(of dimension* $m \times n$*) is multiplied on the right by a diagonal matrix* $\{d_1, d_2, \ldots, d_n\}$*, then the columns of* A *are multiplied by* $d_1, d_2, \ldots, d_n$*, respectively.*

Proof. In fact

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}d_1 & a_{12}d_2 & \cdots & a_{1n}d_n \\ a_{21}d_1 & a_{22}d_2 & \cdots & a_{2n}d_n \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}d_1 & a_{m2}d_2 & \cdots & a_{mn}d_n \end{bmatrix} . \diamond$$

Theorem 2.1.4 *If a rectangular matrix* A *(of dimension* $m \times n$*) is multiplied on the left by a diagonal matrix* $\{d_1, d_2, \ldots, d_m\}$*, then the rows of* A *are multiplied by* $d_1, d_2, \ldots, d_m$*, respectively.*

Proof. In fact

$$\begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} d_1 a_{11} & d_1 a_{12} & \cdots & d_1 a_{1n} \\ d_2 a_{21} & d_2 a_{22} & \cdots & d_2 a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ d_m a_{m1} & d_m a_{m2} & \cdots & d_m a_{mn} \end{bmatrix} . \diamond$$

Theorem 2.1.5 *Let* U *be a unit matrix of order* $n \in \mathbb{N}$ *. Then for every square matrix* A *of order* $n \in \mathbb{N}$ *we have*

$$U\,A = A\,U = A\,.$$

Proof. Obvious. ◊

Theorem 2.1.6 *Let* $A\,,B$ *be two rectangular matrices and* $\alpha \in \mathbb{R}$ *. We have*

(2.1.19) $$(A + B)^T = A^T + B^T$$

(2.1.20) $$(\alpha\,A)^T = \alpha\,A^T$$

(2.1.21) $$(A\,B)^T = B^T A^T\,.$$

Proof. Obvious. ◊

Definition 2.1.21 Let A any non-zero square matrix of order $n \in \mathbb{N}$, U the unit matrix of order n. We say that A is *invertible* if there exists a square matrix B of order n such that

$$A\,B = B\,A = U\,.$$

If A is invertible, B is called the *inverse matrix of* A and denoted A^{-1}. ◊

Theorem 2.1.7 *Le t A be an invertible matrix of order* $n \in \mathbb{N}$. *The inverse of* A *is unique.*

Proof. Let B, C two inverse matrices of A. Hence $A\,B = B\,A = U$ and $A\,C = C\,A = U$. Hence $C = C\,U = C(A\,B) = (C\,A)B = U\,B = B$. ◊

Theorem 2.1.8 *Let* $A\,, B$ *be two invertible matrices of order* $n \in \mathbb{N}$. *The matrix* $A\,B$ *is invertible and we have*

$$(A\,B)^{-1} = B^{-1}A^{-1}\,. \qquad (2.1.22)$$

Proof. We have $(A\,B)(B^{-1}\,A^{-1}) = A(B\,B^{-1})A^{-1} = A\,A^{-1} = U$ and $(B^{-1}\,A^{-1})(A\,B) = B^{-1}(A^{-1}A)B = B^{-1}\,B = U$. From theorem 2.1.7 we