

Chapter 4

Matrices

4.1 Informally

4.1.1 A matrix is a practical way to denote a function defined on the Cartesian product of two finite sets, with values in a certain set X . The advantage of such a table is that the image of an element of the domain can be read directly from the table itself. The position of each element of the table, that is, the row and the column to which it belongs (see below), usually carries an important informational content.

4.1.2 Esempio. The following example consists of three parts, a), b), and c).

- a) Suppose one wants to tabulate the temperature at 8:00 in the morning in 4 cities over the course of a week. Let

$$\mathcal{S} := \{\text{M, Tu, W, Th, F, Sa, Su}\},$$

be the set of the 7 days of the week, and let

$$\mathcal{C} := \{\mathbf{A}(\text{ncona}), \mathbf{B}(\text{ologna}), \mathbf{C}(\text{osenza}), \mathbf{D}(\text{omodossola})\},$$

be the set of the 4 cities. Let moreover

$$T_1 : \mathcal{S} \times \mathcal{C} \longrightarrow \mathbb{R}$$

be the function *temperature at 8:00*, which assigns to each pair

(day of the week, city)

the temperature at 8:00 (on that particular day and in that particular city), expressed in degrees Celsius. The function T_1 can be identified with the rectangular table (6.1) with 4 rows and 7 columns:

$$\begin{array}{ccccccc}
 & M & Tu & W & Th & F & Sa & Su \\
 & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
 A \rightarrow & 10 & 12 & 9 & 8 & 9 & 10 & 12 \\
 B \rightarrow & 10 & 11 & 12 & 10 & 7 & 8 & 13 \\
 C \rightarrow & 15 & 13 & 16 & 13 & 18 & 18 & 17 \\
 D \rightarrow & 8 & 7 & 8 & 10 & 9 & 8 & 11
 \end{array} \tag{4.1}$$

The table shows, for example, that $T_1(Tu, C) = 13$, that is, that the temperature in Cosenza at 8:00 on Tuesday is 13°C . In each column one reads the temperature at 8:00 for every city on the day corresponding to that column itself. In each row one reads the temperature on each day of the week, at 8:00, in the city corresponding to that row itself.

- b) Let now $T_2 : \mathcal{S} \times \mathcal{C} \longrightarrow \mathbb{R}$ be a second measurement (for example: at a different hour or with a different thermometer), given by the table:

$$T_2 := \begin{array}{ccccccc}
 11 & 10 & 8 & 9 & 9 & 11 & 12 \\
 13 & 14 & 11 & 12 & 9 & 11 & 9 \\
 10 & 14 & 15 & 12 & 17 & 19 & 16 \\
 8 & 7 & 9 & 11 & 9 & 10 & 10
 \end{array}.$$

The *arithmetic mean* of the measurements is:

$$T_M = \frac{1}{2}(T_1 + T_2) = \frac{1}{2} \left(\begin{array}{|c|} \hline \begin{array}{cccccc} 10 & 12 & 9 & 8 & 9 & 10 & 12 \\ 10 & 11 & 12 & 10 & 7 & 8 & 13 \\ 15 & 13 & 16 & 13 & 18 & 18 & 17 \\ 8 & 7 & 8 & 10 & 9 & 8 & 11 \end{array} \\ \hline \end{array} + \begin{array}{|c|} \hline \begin{array}{cccccc} 11 & 10 & 8 & 9 & 9 & 11 & 12 \\ 13 & 14 & 11 & 12 & 9 & 11 & 9 \\ 10 & 14 & 15 & 12 & 17 & 19 & 16 \\ 8 & 7 & 9 & 11 & 9 & 10 & 10 \end{array} \\ \hline \end{array} \right) =$$

$$= \frac{1}{2} \begin{array}{|c|} \hline \begin{array}{cccccc} 10 + 11 & 12 + 10 & 9 + 8 & 8 + 9 & 9 + 9 & 10 + 11 & 12 + 12 \\ 10 + 13 & 11 + 14 & 12 + 11 & 10 + 12 & 7 + 9 & 8 + 11 & 13 + 9 \\ 15 + 10 & 13 + 14 & 16 + 15 & 13 + 12 & 18 + 17 & 18 + 19 & 17 + 16 \\ 8 + 8 & 7 + 7 & 8 + 9 & 10 + 11 & 9 + 9 & 8 + 10 & 11 + 10 \end{array} \\ \hline \end{array}$$

$$= \begin{array}{|c|} \hline \begin{array}{cccccc} \frac{21}{2} & \frac{22}{2} & \frac{17}{2} & \frac{17}{2} & \frac{18}{2} & \frac{21}{2} & \frac{24}{2} \\ \frac{23}{2} & \frac{25}{2} & \frac{23}{2} & \frac{22}{2} & \frac{15}{2} & \frac{19}{2} & \frac{22}{2} \\ \frac{25}{2} & \frac{27}{2} & \frac{31}{2} & \frac{25}{2} & \frac{35}{2} & \frac{37}{2} & \frac{33}{2} \\ \frac{16}{2} & \frac{14}{2} & \frac{17}{2} & \frac{21}{2} & \frac{18}{2} & \frac{18}{2} & \frac{21}{2} \end{array} \\ \hline \end{array} = \begin{array}{|c|} \hline \begin{array}{cccccc} 10.5 & 11 & 8.5 & 8.5 & 9 & 10.5 & 12 \\ 11.5 & 12.5 & 11.5 & 11 & 7.5 & 9.5 & 11 \\ 12.5 & 13.5 & 15.5 & 12.5 & 17.5 & 18.5 & 16.5 \\ 8 & 7 & 8.5 & 10.5 & 9 & 9 & 10.5 \end{array} \\ \hline \end{array}$$

- c) The values of the function T_1 are expressed in degrees Celsius. One Réaumur degree is equal to $5/4$ of a degree Celsius. Therefore, to obtain the table R_1 of the temperatures of the 4 cities during the 7 days of the week expressed in degrees Réaumur, it is

enough to multiply T_1 by $5/4$:

$$R_1 = \frac{5}{4} \cdot T_1 = \frac{5}{4} \cdot \begin{array}{|c|c|c|c|c|c|c|} \hline 10 & 12 & 9 & 8 & 9 & 10 & 12 \\ \hline 10 & 11 & 12 & 10 & 7 & 8 & 13 \\ \hline 15 & 13 & 16 & 13 & 18 & 18 & 17 \\ \hline 8 & 7 & 8 & 10 & 9 & 8 & 11 \\ \hline \end{array} = \begin{array}{|c|c|c|c|c|c|c|} \hline \frac{50}{4} & \frac{60}{4} & \frac{45}{4} & \frac{40}{4} & \frac{45}{4} & \frac{50}{4} & \frac{60}{4} \\ \hline \frac{50}{4} & \frac{55}{4} & \frac{60}{4} & \frac{50}{4} & \frac{35}{4} & \frac{40}{4} & \frac{65}{4} \\ \hline \frac{60}{4} & \frac{65}{4} & \frac{80}{4} & \frac{65}{4} & \frac{90}{4} & \frac{90}{4} & \frac{85}{4} \\ \hline \frac{40}{4} & \frac{35}{4} & \frac{40}{4} & \frac{50}{4} & \frac{45}{4} & \frac{40}{4} & \frac{55}{4} \\ \hline \end{array} =$$

$$= \begin{array}{|c|c|c|c|c|c|c|} \hline 12.5 & 15 & 11.25 & 10 & 11.25 & 12.5 & 15 \\ \hline 11.25 & 13.75 & 15 & 12.5 & 8.75 & 10 & 16.25 \\ \hline 15 & 16.25 & 20 & 16.25 & 22.5 & 22.5 & 21.25 \\ \hline 10 & 8.75 & 10 & 12.5 & 11.25 & 10 & 13.75 \\ \hline \end{array}.$$

4.1.3 Readers, male and female, to whom the spirit of Example 6.1.2 will not have escaped, will likewise have understood that matrix manipulations, which will be introduced in the following paragraphs, are governed by absolutely natural operations (those of a vector space!). The lowering of the level of difficulty in reading the present chapter is, therefore, inevitable, since the hidden meaning of the formalism that we shall not hesitate to introduce and develop has already been revealed in advance.

4.2 First formal definitions

It is natural to begin by defining *tables* of elements of arbitrary sets. So let X be an arbitrary set, for example the set of magnetic stickers for playing on the refrigerator at home (as when one was a child).

4.2.1 Definition. A matrix with m rows and n columns, with entries in X , is a function¹

$$A : \underline{m} \times \underline{n} \longrightarrow X.$$

¹Recall that $\underline{m} := \{1, 2, \dots, m\}$ and $\underline{n} = \{1, 2, \dots, n\}$, the sets, respectively, of the first m and the first n natural numbers.

The set of matrices with m rows and n columns will be denoted by $X^{m \times n}$ instead of in the, certainly more precise, way $X^{\underline{m} \times \underline{n}}$. For every $(i, j) \in \underline{m} \times \underline{n}$, $A(i, j)$ is an element of the set X . Since A is a function defined on a finite set of elements, namely on the mn pairs $(i, j) \in \mathbb{Z}^2$ such that $1 \leq i \leq m$ and $1 \leq j \leq n$, the function A can be represented by listing its images, that is, by explicitly listing the *entries* of the table:

$$A := \begin{pmatrix} A(1, 1) & A(1, 2) & \dots & A(1, n) \\ A(2, 1) & A(2, 2) & \dots & A(2, n) \\ \vdots & \vdots & \ddots & \vdots \\ A(m, 1) & A(m, 2) & \dots & A(m, n) \end{pmatrix}.$$

Moreover, if the image of the element $(i, j) \in \underline{m} \times \underline{n}$ is a_{ij} , that is, if $a_{ij} = A(i, j)$, it may be written, as usual, in the form

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \quad (4.2)$$

If, on the other hand, one does not wish to specify the image of the function A at the element $(i, j) \in \underline{m} \times \underline{n}$, one will simply write $A(i, j)$. In our course we shall interpret the “table” on the right-hand side of (6.2) precisely as a function, in the following way:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} (i, j) = a_{ij},$$

where the value of the function at (i, j) is precisely the element a_{ij} of the set X , that is, the *entry* of the matrix located in the i -th row and the j -th column.

4.2.2 A matrix with one row and n columns, that is,

$$R : \{1\} \times \underline{n} \rightarrow X,$$

is called a *row matrix* or, briefly, a *row* (with n columns). A matrix

$$C : \underline{m} \times \{1\} \rightarrow X,$$

with one column and m rows is called a *column matrix* or, briefly, a *column* (with m rows). When $m = n$, that is, when the number of rows is equal to the number of

columns, the matrix is said to be *square*. Consistently, the set of all square $n \times n$ matrices will be denoted by $X^{n \times n}$ (instead of in the more precise but also more pedantic way $X^{\mathfrak{n} \times \mathfrak{n}}$).

4.2.3 Esempio. Let $X = \{\heartsuit, \spadesuit, \diamondsuit, \clubsuit\}$. Then

$$A := \begin{pmatrix} \heartsuit & \spadesuit & \heartsuit \\ \heartsuit & \diamondsuit & \clubsuit \\ \diamondsuit & \diamondsuit & \spadesuit \end{pmatrix}$$

is a square 3×3 matrix. For example $A(2, 3) = \clubsuit$. What is $A(3, 2)$?

Once the set X has been arbitrarily fixed, it is important to establish when two matrices are equal. This is determined by the general criterion of equality between functions, very simple but of crucial importance.

Two $m \times n$ matrices are equal if and only if they are equal as functions from $\underline{m} \times \underline{n} \rightarrow X$.

More explicitly:

4.2.4 Definition. Let $A, B \in X^{\mathfrak{m} \times \mathfrak{n}}$. Then

$$A = B \iff A(i, j) = B(i, j), \quad \forall (i, j) \in \underline{m} \times \underline{n}. \quad (4.3)$$

In other words, two matrices are equal if and only if they have the same entries.

4.2.5 Let $A \in X^{\mathfrak{m} \times \mathfrak{n}}$. For every $1 \leq j \leq n$, the j -th column of A is the function

$$C_j(A) : \underline{m} \longrightarrow X,$$

defined by:

$$C_j(A)(i) = A(i, j), \quad \forall 1 \leq i \leq m \quad (4.4)$$

In other words:

$$C_j(A) = \begin{pmatrix} A(1, j) \\ \vdots \\ A(m, j) \end{pmatrix}.$$

The i -th row of the column $C_j(A)$ is $C_j(A)(i)$, that is, the i -th component of the column $C_j(A)$.

4.2.6 Similarly, if $1 \leq i \leq m$, the i -th row of the matrix A is the function:

$$R_i(A) : \underline{n} \longrightarrow X,$$

defined by

$$R_i(A)(j) = A(i, j), \quad \forall 1 \leq j \leq n. \quad (4.5)$$

In other words:

$$R_i(A) = (A(i, 1), \dots, A(i, n)).$$

The j -th column of the row $R_i(A)$ is $R_i(A)(j)$, that is, the j -th entry of $R_i(A)$.

$$\begin{array}{c}
 C_j \\
 \begin{array}{|c|} \hline a_{1j} \\ \hline \vdots \\ \hline \end{array} \\
 \begin{array}{|c|} \hline a_{i1} \quad \dots \quad a_{ij} \quad \dots \quad a_{in} \\ \hline \end{array} \\
 \begin{array}{|c|} \hline \vdots \\ \hline a_{mj} \\ \hline \end{array}
 \end{array}$$

Summing up:

$$R_i(C_j(A)) = C_j(R_i(A)) = A(i, j) \quad (4.6)$$

4.2.7 Warning. In the definition of row and column there is a strong abuse of notation. Strictly speaking, a column is a matrix with m rows and one column and, therefore, it should have been defined as a function $\underline{m} \times \{1\} \rightarrow X$ such that

$$C_j(A)(i, 1) = A(i, j). \quad (4.7)$$

Similarly, a row should be a row matrix, that is, a function $\{1\} \times \underline{n} \rightarrow X$ defined by

$$R_i(A)(1, j) = A(i, j). \quad (4.8)$$

One has chosen, in order not to burden the notation unreasonably, to use the letter C for columns and the letter R for rows. This will indicate when one is referring to rows or to columns, which was indeed the only purpose of the position of the 1 in the expressions (6.7) and (6.8)

4.2.8 It is natural to observe that every $A \in X^{\mathfrak{m} \times \mathfrak{n}}$ can be written, equivalently, as a *column of m rows* or as a *row of n columns*:

$$A := \begin{pmatrix} R_1(A) \\ \vdots \\ R_m(A) \end{pmatrix} = (C_1(A), \dots, C_n(A)).$$

4.2.9 Esempio. Consider the matrix A of example 6.2.3. Then:

$$R_1(A) = (\heartsuit \quad \spadesuit \quad \heartsuit) \quad \text{and} \quad C_3(A) = \begin{pmatrix} \heartsuit \\ \clubsuit \\ \spadesuit \end{pmatrix}$$

■

4.2.10 Proposition. *Two matrices are equal if and only if they have all columns equal, or all rows equal.*

Proof. Let $A = B$. Then for every $(i, j) \in \mathfrak{m} \times \mathfrak{n}$ one has $A(i, j) = B(i, j)$, by Definition 6.2.4 of equality of matrices, whence, for every $1 \leq i \leq m$:

$$R_i(A)(j) = A(i, j) = B(i, j) = R_i(B)(j)$$

for every $1 \leq j \leq n$, that is, $R_i(A) = R_i(B)$, again by criterion 6.2.4 of equality between matrices. Similarly, for every $1 \leq j \leq n$ one has

$$C_j(A)(i) = A(i, j) = B(i, j) = C_j(B)(i), \tag{4.9}$$

for every $1 \leq i \leq m$, that is, $C_j(A)$ and $C_j(B)$ are equal, again by 6.2.4. Conversely, suppose that all rows of A and B are equal, that is, that for every $1 \leq i \leq m$ one has

$$R_i(A) = R_i(B).$$

Then, for every $(i, j) \in \mathfrak{m} \times \mathfrak{n}$:

$$A(i, j) = R_i(A)(j) = R_i(B)(j) = B(i, j),$$

where one has used definition (6.5) of the $(i$ -th) row of a matrix. Therefore $A = B$, since $A(i, j) = B(i, j)$ for every $(i, j) \in \underline{m} \times \underline{n}$ and by criterion (6.2.4). Entirely similar is the proof starting from the equality of the columns. If $C_j(A) = C_j(B)$ for every $1 \leq j \leq n$, then, for every $(i, j) \in \underline{m} \times \underline{n}$, one has:

$$A(i, j) = C_j(A)(i) = C_j(B)(i) = B(i, j),$$

where one has used Definition (6.9) of the $(j$ -th) column of a matrix. Therefore $A = B$, since $A(i, j) = B(i, j)$ for every $(i, j) \in \underline{m} \times \underline{n}$ and by criterion (6.2.4). ■

4.2.11 Transposition of rows and columns. There exists a natural bijection

$$^T : X^{1 \times \underline{n}} \longrightarrow X^{\underline{n} \times 1}$$

between row matrices with n columns, $X^{1 \times \underline{n}}$, and column matrices with n rows, $X^{\underline{n} \times 1}$, called *row transposition*. It is defined as follows: for every row $R \in X^{1 \times \underline{n}}$, R^T is the unique column defined by:

$$R^T(i) = R(i),$$

for every $1 \leq i \leq n$, that is, such that the i -th row of the column R^T coincides with the i -th column of the row R . By abuse of notation one will denote by the same notation T the function

$$^T : X^{\underline{n} \times 1} \longrightarrow X^{1 \times \underline{n}}$$

called *column transposition*, defined by

$$C^T(j) = C(j),$$

for every $1 \leq j \leq n$. For every $R \in X^{1 \times \underline{n}}$ and for every $C \in X^{\underline{n} \times 1}$, define, respectively, in order to streamline the notation:

$$R^{TT} = (R^T)^T \quad \text{and} \quad C^{TT} = (C^T)^T.$$

4.2.12 The two types of transposition are clearly bijections (see more directly the following paragraph 6.2.13). Indeed, the equality $R_1^T = R_2^T$ implies:

$$R_1^T(j) = R_2^T(j)$$

equivalent to:

$$R_1(j) = R_2(j)$$

whence $R_1 = R_2$, by equality of all entries. Moreover, for every $R \in X^{1 \times n}$ and for every $1 \leq j \leq n$ one has:

$$R^{TT}(j) = ((R^T)^T)(j) = R^T(j) = R(j)$$

that is,

$$R^{TT} := (R^T)^T = R, \quad (4.10)$$

since the two matrices R^{TT} and R have the same columns. Similarly

$$C^{TT}(i) = (C^T)^T(i) = C^T(i) = C(i)$$

and this proves that $C^{TT} = C$. Therefore T is also surjective. Indeed, for every $C \in X^{n \times 1}$, $C^T \in X^{1 \times n}$ and, moreover,

$$C^{TT} := (C^T)^T = C. \quad (4.11)$$

Thus one has found a row, namely C^T , whose transpose is C . In the same way one proves that $^T : X^{n \times 1} \rightarrow X^{1 \times n}$ is injective and surjective. ■

4.2.13 Behind the formalism. Although the definition and the properties of transposition, mentioned in paragraph 6.2.12, may appear complicated, they are nothing other than the formal translation of a very simple fact: every ordered n -tuple of elements of X may be written either as a row or as a column, and every row may be rewritten as a column and every column may be rewritten as a row. Thus, to be more explicit, and this is really the idea that each of us should have in mind, row transposition is the function:

$$\left\{ \begin{array}{lcl} ^T : & X^{1 \times n} & \longrightarrow X^{n \times 1} \\ & (x_1, \dots, x_n) & \longmapsto (x_1, \dots, x_n)^T := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{array} \right.$$

Similarly

$$\left\{ \begin{array}{lcl} ^T : & X^{n \times 1} & \longrightarrow X^{1 \times n} \\ & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^T = (x_1, \dots, x_n) \end{array} \right.$$

It is therefore clear that, after rewriting the elements of a row as a column and then laying them out again along a row, one recovers the same starting row (that is, as proved formally: $R^{TT} = R$). The same holds for columns. Lay it down and raise it up again, and it will be the same column.

4.3 Transpose of a matrix

Let $A \in X^{\mathfrak{m} \times \mathfrak{n}}$ be a matrix with m rows and n columns.

4.3.1 *The transpose of A is the unique matrix $A^T \in X^{\mathfrak{n} \times \mathfrak{m}}$ defined by:*

$$R_j(A^T) = C_j(A)^T. \quad (4.12)$$

In other words, A^T is the matrix whose *rows* are the transposes of the *columns* of A . More explicitly, writing A as the row of its n columns:

$$A = (C_1(A), C_2(A), \dots, C_n(A))$$

the transpose is:

$$A^T = \begin{pmatrix} C_1(A)^T \\ \vdots \\ C_n(A)^T \end{pmatrix}.$$

4.3.2 Proposition. *Let $A \in X^{\mathfrak{m} \times \mathfrak{n}}$. Then, for every $i \in \mathfrak{n}$*

$$C_i(A^T) = R_i(A)^T. \quad (4.13)$$

Proof.

$$C_i(A^T)^T = R_i(A^{TT})$$

applying (6.12) to A^T . Therefore $C_i(A^T)^T = R_i(A)$. Taking the transpose of both sides, one obtains $C_i(A^T) = R_i(A)^T$ as was to be proved. ■

Had one adopted (6.13) as the definition of the transpose of A , one would have proved (6.12) as a consequence. In other words, (6.13) is a definition of the transpose of a matrix equivalent to (6.12). Writing A in the form

$$A = \begin{pmatrix} R_1(A) \\ \vdots \\ R_m(A) \end{pmatrix}$$

its transpose is:

$$A^T = (R_1(A)^T, \dots, R_m(A)^T).$$

4.3.3 Proposition. *The function:*

$$^T : X^{\mathfrak{m} \times \mathfrak{n}} \longrightarrow X^{\mathfrak{n} \times \mathfrak{m}}$$

is a bijection. Moreover $^{TT} : X^{\mathfrak{m} \times \mathfrak{n}} \longrightarrow X^{\mathfrak{m} \times \mathfrak{n}}$ is the identity.

Proof. Let us first show that $A^{TT} = (A^T)^T = A$ for every $A \in X^{\mathfrak{m} \times \mathfrak{n}}$. It is sufficient to show, for example, that A^{TT} and A have all columns equal.

$$\begin{aligned}
 C_j(A^{TT}) &= C_j((A^T)^T) = && \text{(Definition of double transposition)} \\
 &= R_j(A^T)^T = && \text{(equality (6.13))} \\
 &= (C_j(A)^T)^T = && \text{(equality (6.13))} \\
 &= C_j(A)^{TT} = && \text{(Definition of double transposition)} \\
 &= C_j(A) && \text{(equality (6.11) applied to } C_j(A)\text{).}
 \end{aligned}$$

Therefore $A^{TT} = A$ for every A , by Proposition 6.2.10. At this point one has that T is injective. Indeed $A^T = B^T$ implies $(A^T)^T = (B^T)^T$, that is $A^{TT} = B^{TT}$, that is $A = B$. Finally, if $B \in X^{\mathfrak{n} \times \mathfrak{m}}$, and if $A = B^T \in X^{\mathfrak{m} \times \mathfrak{n}}$, then $A^T = (B^T)^T = B^{TT} = B$, and therefore T is surjective, that is, $^T : X^{\mathfrak{m} \times \mathfrak{n}} \rightarrow X^{\mathfrak{n} \times \mathfrak{m}}$ is bijective. ■

4.3.4 Proposition. *Let $A \in X^{\mathfrak{n} \times \mathfrak{n}}$. If $A^T = A$, then A is a square matrix.*

Indeed, if $A = A^T$, then $A \in X^{\mathfrak{n} \times \mathfrak{n}}$ and $A \in X^{\mathfrak{n} \times \mathfrak{m}}$, that is, $m = n$. ■

Clearly the converse is false. Indeed, if $A \in \{a, b, c\}^{2 \times 2}$ is given by:

$$A := \begin{pmatrix} a & b \\ c & a \end{pmatrix}.$$

then A is certainly square, but

$$A^T = \begin{pmatrix} a & c \\ b & a \end{pmatrix} \neq \begin{pmatrix} a & b \\ c & a \end{pmatrix} = A.$$

4.3.5 Definition. *A square matrix $A \in X^{\mathfrak{n} \times \mathfrak{n}}$ is said to be symmetric if and only if $A = A^T$.*

All 1×1 matrices are symmetric. Indeed they have only one row and one column. The set $X^{1 \times 1}$ can be identified with the set X itself, in the sense that there exists a natural bijection $X^{1 \times 1} \rightarrow X$, defined by $(x) \mapsto x$. It is the bijection that *removes the parentheses*. Formally:

$$R_1((x)^T) = (C_1(x))^T = (R_1(x)^T)^T = R_1((x)),$$

whence $(x)^T = (x)$.

4.4 Vector spaces of matrices

In the previous sections of this chapter matrices with entries in a set without particular structures were introduced. If the set X has further structures (topological, algebraic, differential), these are reflected on the set $X^{\mathfrak{m} \times \mathfrak{n}}$.

4.4.1 Here is, so as not to lose the readers, male and female, or in order to recover them, a quick recall of Paragraph ???. As usual, $\mathbb{K}$ will denote a subfield of the complex field (for example $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$). Let $\mathbb{K}^{m \times n}$ be the set of all $m \times n$ matrices with values in $\mathbb{K}$. As anticipated in Section ??, the set $\mathbb{K}^{m \times n}$ is a $\mathbb{K}$ -vector space. The linear combination of two matrices is as in ??, and is a particular case of the notion of linear combination of scalar-valued functions. The vector space $\mathbb{K}^{n \times 1}$ of *column matrices*, or *columns*, will be denoted, for brevity, by $\mathbb{K}^n$. Similarly the vector space $\mathbb{K}^{1 \times n}$ of row matrices, or *rows*, will be denoted by $(\mathbb{K}^n)^\vee$. If $R \in (\mathbb{K}^n)^\vee$ one will write $R(j)$ instead of $R(1, j)$, and if $C \in \mathbb{K}^n$ one will write briefly $C(i)$ instead of $C(i, 1)$. As a particular case of (??), one has that the linear combination of two rows $R_1, R_2 \in (\mathbb{K}^n)^\vee$, with coefficients $\lambda, \mu \in \mathbb{K}$, is the unique row $\lambda R_1 + \mu R_2$ such that

$$(\lambda R_1 + \mu R_2)(j) = \lambda R_1(j) + \mu R_2(j), \quad (4.14)$$

for every $1 \leq j \leq n$, just as the linear combination of two columns $C_1, C_2 \in \mathbb{K}^n$ with coefficients $\lambda, \mu \in \mathbb{K}$ is the unique column $\lambda C_1 + \mu C_2$ such that

$$(\lambda C_1 + \mu C_2)(i) = \lambda C_1(i) + \mu C_2(i), \quad (4.15)$$

for every $1 \leq i \leq n$. The *transposition* $(\mathbb{K}^n)^\vee \rightarrow \mathbb{K}^n$ is linear, in the sense that

$$(\lambda \alpha + \mu \beta)^T = \lambda \alpha^T + \mu \beta^T$$

for every $\alpha, \beta \in (\mathbb{K}^n)^\vee$ and every $\lambda, \mu \in \mathbb{K}$, an obvious verification left to the reader.

4.4.2 One will furthermore observe that the functions $C_j : \mathbb{K}^{m \times n} \rightarrow \mathbb{K}^n$ and $R_i : \mathbb{K}^{m \times n} \rightarrow (\mathbb{K}^n)^\vee$ are linear in the sense that the j -th column of the linear combination of two matrices is the linear combination of the j -th columns of the matrices, and that the i -th row of the linear combination of two matrices is the linear combination of the i -th rows of the two matrices. Formally, for every $1 \leq i \leq m$ and every $1 \leq j \leq n$:

$$R_i(\lambda A + \mu B) = \lambda R_i(A) + \mu R_i(B) \quad \text{and} \quad C_j(\lambda A + \mu B) = \lambda C_j(A) + \mu C_j(B) \quad (4.16)$$

4.5 Transposition and conjugation

Recall that the transpose matrix A^T of $A \in \mathbb{K}^{m \times n}$ is defined equivalently by

$$C_i(A^T) = R_i(A)^T \quad \text{or} \quad R_j(A^T) = C_j(A)^T,$$

so that, in particular, the transpose of a 1×1 matrix coincides with itself (it is trivially symmetric). Therefore

$$A^T(i, j) = C_j(R_i(A^T)) = C_j((C_i(A))^T) = (R_j(C_i(A)))^T = R_j(C_i(A)) = A(j, i), \quad (4.17)$$

expressing the fact that the (i, j) entry of the transpose A^T of A coincides with the (j, i) entry of the matrix A itself. One may therefore easily see that:

4.5.1 Proposition. *The map*

$$\begin{cases} T : \mathbb{K}^{m \times n} & \longrightarrow & \mathbb{K}^{n \times m} \\ A & \longmapsto & A^T \end{cases} \quad (4.18)$$

is linear, that is, for every $A, B \in \mathbb{K}^{m \times n}$ and every $\lambda, \mu \in \mathbb{K}$:

$$(\lambda A + \mu B)^T = \lambda A^T + \mu B^T. \quad (4.19)$$

Moreover $A^{TT} := (A^T)^T = A$, for every $A \in \mathbb{K}^{m \times n}$.

Proof. It is enough to verify equality between the entries of the left-hand side and the right-hand side of (6.18). For every $1 \leq j \leq n$:

$$(\lambda A + \mu B)^T(i, j) = (\lambda A + \mu B)(j, i) =$$

and, by the definition of *linear combination of matrices* (??), one has:

$$= \lambda A(j, i) + \mu B(j, i) = \lambda A^T(i, j) + \mu B^T(i, j) = (\lambda A^T + \mu B^T)(i, j)$$

where in the first equality one used (6.17), and in the second equality the definition of linear combination of matrices. ■

4.5.2 since, as assumed, $\mathbb{K}$ is a subfield of the complex field $\mathbb{C}$, there exists a function

$$\begin{cases} \mathbb{K} & \longrightarrow & \mathbb{K} \\ \lambda & \longmapsto & \bar{\lambda} \end{cases},$$

called *conjugation*, defined, as is well known, by: if $z := a + bi \in \mathbb{C}$, then $\bar{z} = a - bi$. Conjugation is clearly the identity for all complex numbers with zero imaginary

part, in particular for real numbers. Recall that conjugation satisfies the following properties, easily checked:

$$\overline{\lambda + \mu} = \bar{\lambda} + \bar{\mu} \quad \text{and} \quad \overline{\lambda\mu} = \bar{\lambda}\bar{\mu}. \quad (4.20)$$

Moreover the modulus of a scalar $\lambda \in \mathbb{K}$ is $|\lambda| := \sqrt{\lambda\bar{\lambda}}$, which of course coincides with the usual modulus of a real number, if $\lambda \in \mathbb{R}$.

Conjugation extends to the space of matrices $\mathbb{K}^{m \times n}$. If $A \in \mathbb{K}^{m \times n}$, one defines the *conjugate matrix* through the relation:

$$\overline{A}(i, j) = \overline{A(i, j)}$$

that is, the conjugate of a matrix is the matrix whose entries are the conjugates of the entries. The conjugation operation on matrices is not necessarily linear. Indeed:

$$\begin{aligned} \overline{(\lambda A + \mu B)}(i, j) &= \overline{(\lambda A + \mu B)(i, j)} = && \text{(definition of conjugation)} \\ &= \overline{(\lambda A(i, j) + \mu B(i, j))} = && \text{(definition of linear combination of matrices)} \\ &= \bar{\lambda} \overline{A(i, j)} + \bar{\mu} \overline{B(i, j)} = && \text{(conjugation of sum and product of complex numbers)} \\ &= \bar{\lambda} \overline{A}(i, j) + \bar{\mu} \overline{B}(i, j) = && \text{(definition of conjugation)} \\ &= (\bar{\lambda} \overline{A} + \bar{\mu} \overline{B})(i, j) = && \text{(definition of linear combination of matrices)} \end{aligned}$$

whence, finally:

$$\overline{(\lambda A + \mu B)}(i, j) = \bar{\lambda} \overline{A}(i, j) + \bar{\mu} \overline{B}(i, j).$$

The conjugation operation is linear if and only if $\mathbb{K}$ is a subfield of $\mathbb{R}$, when the latter is viewed as a subfield of the complex numbers with zero imaginary part. Indeed, in this case conjugation is the identity. It is instead *antilinear* if $\mathbb{K} = \mathbb{C}$, in the sense that the scalars occurring in the linear combination of the conjugate matrices are themselves conjugated.

4.5.3 Proposition *The operations of conjugation and transposition commute, that is:*

$$\overline{A}^T = \overline{A^T} \quad (4.21)$$

Proof. This is one of those proofs that ought to be left as an exercise. It is enough, however, to observe that for every $(i, j) \in \underline{m} \times \underline{n}$:

$$\overline{A}^T(i, j) = \overline{A}(j, i) = \overline{A(j, i)} = \overline{A^T(i, j)} = \overline{A^T}(i, j)$$

and the reader will be left the exercise of explaining why each equality holds, specifying which property or definition has been invoked each time.

4.6 Product of matrices

4.6.1 Let m, n, p be arbitrary positive integers. The purpose of this and the next section is to define a certain bilinear map:

$$\begin{cases} \mathbb{K}^{m \times p} \times \mathbb{K}^{p \times n} & \longrightarrow \mathbb{K}^{m \times n}, \\ (A, B) & \longmapsto A \cdot B. \end{cases}$$

which associates with an ordered pair of matrices respectively $m \times p$ and $p \times n$ a third matrix, denoted by $A \cdot B$. The matrix $A \cdot B \in \mathbb{K}^{m \times n}$ will be called the *product* of the matrices A and B , and it will be defined, as prescribed, only if the number of columns of A coincides with the number of rows of B . The *bilinearity* of matrix multiplication means that for every $A, A_1, A_2 \in \mathbb{K}^{m \times p}$ and every $B, B_1, B_2 \in \mathbb{K}^{p \times n}$ and every $\lambda, \mu \in \mathbb{K}$ one has

$$(\lambda A_1 + \mu A_2)B = \lambda A_1 \cdot B + \mu A_2 \cdot B, \quad (4.22)$$

which expresses *linearity with respect to the first argument*, and

$$A(\lambda B_1 + \mu B_2) = \lambda A \cdot B_1 + \mu A \cdot B_2. \quad (4.23)$$

which expresses *linearity with respect to the second argument*. The product of matrices will in general be defined through a concatenation of special cases.

4.6.2 Definition. The product of $A \in \mathbb{K}^{m \times n}$ by the column $\mathbf{u} \in \mathbb{K}^n$ is the column:

$$A\mathbf{u} = \mathbf{u}(1)C_1(A) + \dots + \mathbf{u}(n)C_n(A) = \sum_{i=1}^n \mathbf{u}(i)C_i(A). \quad (4.24)$$

In other words the product $A\mathbf{u}$ is the linear combination of the columns of A , where the coefficient of the j -th column of A is precisely the j -th entry of the vector $\mathbf{u}$.

4.6.3 Esempio. Let

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 4 & 1 & 2 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{u} = \begin{pmatrix} 2 \\ 3 \\ -2 \\ 5 \end{pmatrix}.$$

Then

$$\begin{aligned} A \cdot \mathbf{u} &= \begin{pmatrix} 1 & 2 & 3 & 1 \\ 4 & 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \\ -2 \\ 5 \end{pmatrix} = 2C_1(A) + 3C_2(A) - 2C_3(A) + 5C_4(A) = \\ &= 2 \begin{pmatrix} 1 \\ 4 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} 3 \\ 2 \end{pmatrix} + 5 \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \\ &= \begin{pmatrix} 2 \\ 8 \end{pmatrix} + \begin{pmatrix} 6 \\ 3 \end{pmatrix} + \begin{pmatrix} -6 \\ -4 \end{pmatrix} + \begin{pmatrix} 5 \\ 15 \end{pmatrix} = \begin{pmatrix} 2+6-6+5 \\ 8+3-4+15 \end{pmatrix} = \begin{pmatrix} 7 \\ 22 \end{pmatrix}. \end{aligned}$$

Easy, isn't it?

4.6.4 Proposition. *The map*

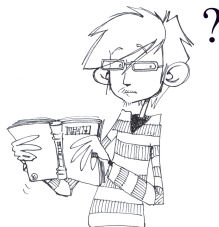
$$\begin{cases} \mathbb{K}^{m \times n} \times \mathbb{K}^n & \longrightarrow \mathbb{K}^m \\ (A, \mathbf{u}) & \mapsto A\mathbf{u}, \end{cases} \quad (4.25)$$

is bilinear, that is, for every $A, B \in \mathbb{K}^{m \times n}$, every $\mathbf{u}, \mathbf{v} \in \mathbb{K}^n$ and every $\lambda, \mu \in \mathbb{K}$:

$$(\lambda A + \mu B)\mathbf{u} = \lambda \cdot A\mathbf{u} + \mu \cdot B\mathbf{u}, \quad (4.26)$$

and

$$A(\lambda \mathbf{u} + \mu \mathbf{v}) = \lambda A \cdot \mathbf{u} + \mu \cdot A\mathbf{v}. \quad (4.27)$$



Proof. As for (6.26), one has:

$$\begin{aligned}
(\lambda A + \mu B)\mathbf{u} &= \sum_{i=1}^n \mathbf{u}(j)C_j(\lambda A + \mu B) = && \text{(definition (6.24) of} \\
&&& \text{the product of a ma-} \\
&&& \text{trix by a column)} \\
&= \sum_{i=1}^n \mathbf{u}(j)(\lambda C_j(A) + \mu C_j(B)) = && \text{(applying the second} \\
&&& \text{of (6.16))} \\
&= \sum_{i=1}^n \lambda \mathbf{u}(j)C_j(A) + \mu \mathbf{u}(j)C_j(B) = && \text{(properties of sums and} \\
&&& \text{products of scalars)} \\
&= \lambda \sum_{i=1}^n \mathbf{u}(j)C_j(A) + \mu \sum_{j=1}^n \mathbf{u}(j)C_j(B) = && (\lambda \text{ and } \mu \text{ factored out)} \\
&= \lambda A\mathbf{u} + \mu B\mathbf{u}. && \text{(definition (6.24) of} \\
&&& \text{the product of a ma-} \\
&&& \text{trix by a column)}
\end{aligned}$$

Equation (6.27) is proved similarly:

$$\begin{aligned}
A(\lambda \mathbf{u} + \mu \mathbf{v}) &= \sum_{j=1}^n (\lambda \mathbf{u} + \mu \mathbf{v})(j)C_j(A) = && \text{(definition (6.24) of} \\
&&& \text{the product of a ma-} \\
&&& \text{trix by a column)} \\
&= \sum_{j=1}^n (\lambda \mathbf{u}(j) + \mu \mathbf{v}(j))C_j(A) = && \text{(definition (6.15)} \\
&&& \text{of linear combination} \\
&&& \text{of column vectors)} \\
&= \sum_{j=1}^n \lambda \mathbf{u}(j)C_j(A) + \sum_{j=1}^n \mu \mathbf{v}(j)C_j(A) = && \text{(distributive property} \\
&&& \text{of addition with respect to} \\
&&& \text{multiplication in a field)} \\
&= \lambda \sum_{j=1}^n \mathbf{u}(j)C_j(A) + \mu \sum_{j=1}^n \mathbf{v}(j)C_j(A) = && (\lambda \text{ and } \mu \text{ factored out)} \\
&= \lambda \cdot A\mathbf{u} + \mu \cdot A\mathbf{v} && \text{(definition (6.24) of} \\
&&& \text{the product of a ma-} \\
&&& \text{trix by a column)}
\end{aligned}$$

■

4.6.5 As a consequence of definition (6.24), one has:

$$A \cdot \mathbf{e}_j = C_j(A), \quad (4.28)$$

that is, multiplication of A by the j -th coordinate column $\mathbf{e}_j$ yields the j -th column of the matrix A . Moreover, if

$$\alpha = (a_1, \dots, a_n) = \sum_{i=1}^n a_i \boldsymbol{\epsilon}_i \in (\mathbb{K}^n)^\vee \quad \text{and} \quad \mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \sum_{j=1}^n u_j \mathbf{e}_j \in \mathbb{K}^n.$$

are respectively an arbitrary row and an arbitrary column:

$$\alpha \cdot \mathbf{u} = \sum_{i=1}^n \alpha(i) \mathbf{u}(i) = u_1 a_1 + \dots + u_n a_n \quad (4.29)$$

Thus, in particular, if C is a column and $\boldsymbol{\epsilon}_i$ is the i -th coordinate row:

$$\boldsymbol{\epsilon}_j \cdot C = \boldsymbol{\epsilon}_j(1)C(1) + \boldsymbol{\epsilon}_j(i) \cdot C(i) + \dots + \boldsymbol{\epsilon}_j(n)C(n) \quad (4.30)$$

Since $\boldsymbol{\epsilon}_j(i) = \delta_{ij}$, in the right-hand side of (6.30) all terms are zero except the j -th one, that is:

$$\boldsymbol{\epsilon}_j \cdot C = \sum_{j=1}^n \boldsymbol{\epsilon}_j(j)C(j) = \sum_{j=1}^n \delta_{jj}C(j) = 1 \cdot C(j) = C(j).$$

4.6.6 The i -th component of the product (6.24) of a matrix by a column is given by:

$$\begin{aligned} (A\mathbf{u})(i) &= (\sum_{j=1}^n \mathbf{u}(j) \cdot C_j(A))(i) = && \text{(definition (6.24) of the product} \\ &&& \text{of a matrix by a column)} \\ &= \sum_{j=1}^n \mathbf{u}(j) \cdot C_j(A)(i) = && \text{(definition (6.15) of linear} \\ &&& \text{combination of column vectors)} \\ &= \sum_{j=1}^n \mathbf{u}(j) \cdot A(i, j) = && \text{(definition of column of } A) \\ &= \sum_{j=1}^n \mathbf{u}(j) \cdot R_i(A)(j), && \text{(definition of row of } A) \\ &= R_i(A)\mathbf{u}, && \text{(definition (6.30) of row by} \\ &&& \text{column).} \end{aligned}$$

whence the useful expression:

$$A\mathbf{u} = \begin{pmatrix} R_1(A) \cdot \mathbf{u} \\ \vdots \\ R_m(A) \cdot \mathbf{u} \end{pmatrix}. \quad (4.31)$$

And finally one is in a position to define in full generality the product of an arbitrary pair of matrices $(A, B) \in \mathbb{K}^{m \times p} \times \mathbb{K}^{p \times n}$.

4.6.7 Definition. *The product $A \cdot B \in \mathbb{K}^{m \times n}$ is the unique matrix such that*

$$(A \cdot B)\mathbf{u} = A \cdot (B \cdot \mathbf{u}). \quad (4.32)$$

From the definition it follows that the (i, j) entry of the product $A \cdot B$ is given by:

$$\begin{aligned} (A \cdot B)(i, j) &= R_i(C_j(A \cdot B)) && \text{(formula (6.6))} \\ &= R_i((A \cdot B)\mathbf{e}_j) = && \text{(formula (6.28) for the } j\text{-th} \\ &&& \text{column of } A \cdot B) \\ &= R_i(A(B \cdot \mathbf{e}_j)) = && \text{(Definition 6.6.7 of the product of} \\ &&& \text{matrices)} \\ &= R_i(A \cdot C_j(B)). && \text{(again formula (6.28)).} \end{aligned}$$

Since $R_i(A \cdot C_j(B))$ is the i -th component of the column $A \cdot C_j(B)$, by (6.31) it coincides with the product of the row $R_i(A)$ by the column $C_j(B)$, that is:

$$(A \cdot B)(i, j) = R_i(A) \cdot C_j(B). \quad (4.33)$$

In other words

The matrix $A \cdot B$ is the matrix whose entry in position (i, j) is obtained by multiplying the i -th row of A by the j -th column of B according to (6.29).

4.6.8 Here are some consequences of the definition of matrix multiplication. First of all note that

$$C_j(A \cdot B) = (AB) \cdot \mathbf{e}_j = A(B\mathbf{e}_j) = AC_j(B),$$

whence:

$$A \cdot B = A(C_1(B), C_2(B), \dots, C_n(B)) = (A \cdot C_1(B), \dots, A \cdot C_n(B)), \quad (4.34)$$

that is,

the product $A \cdot B$ is the matrix whose j -th column is the product of the matrix A by the j -th column $C_j(B)$ of the matrix B .

Secondly, for every $A \in \mathbb{K}^{m \times n}$ and every $1 \leq i \leq m$ and every $1 \leq j \leq n$ one has

$$\epsilon_i A e_j := (\epsilon_i A) e_j = \epsilon_i \cdot (A e_j) = A(i, j) \quad (4.35)$$

where, as usual, ϵ_i is the i -th coordinate row, whose entries are all zero except the i -th, which is equal to 1. The first equality in (6.35) is a special case of Definition 6.6.7. Moreover:

$$\epsilon_i \cdot (A e_j) = \epsilon_i \cdot C_j(A) = \sum_{h=1}^n \epsilon_i(h) C_j(A)(h) = \sum_{h=1}^n \delta_{ih} C_j(A)(h) = C_j(A)(i) = A(i, j)$$

where in the first equality one used formula (6.28), in the second the definition (6.29) of product of a row by a column, in the third the definition of the i -th coordinate row. As for the fourth equality, all summands in the sum are zero except the one for which $h = i$. This leads to the last equality where one used the fact that the entry $A(i, j)$ of A is the i -th component of the j -th column. In particular one sees that

$$C_j(R_i(A)) = A(i, j) = (\epsilon_i \cdot A) e_j = C_j(\epsilon_i \cdot A)$$

that is, the analogue of (6.28):

$$\epsilon_i A = R_i(A), \quad (4.36)$$

for every $1 \leq i \leq n$.

4.6.9 Esempio. Suppose one wishes to multiply the matrices $A \in \mathbb{K}^{2 \times 4}$ and $B \in \mathbb{K}^{4 \times 3}$, where:

$$A := \begin{pmatrix} 1 & -1 & 1 & 3 \\ 2 & 1 & -3 & 1 \end{pmatrix} \quad e \quad B := \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & -1 \\ 3 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

Method I: rows by columns. One multiplies each row of A , in order, by each column of B :

$$\begin{aligned}
 A \cdot B &= \begin{pmatrix} (1, -1, 1, 3) \cdot \begin{pmatrix} 1 \\ 1 \\ 3 \\ 2 \end{pmatrix} & (1, -1, 1, 3) \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix} & (1, -1, 1, 3) \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \\ (2, 1, -3, 1) \cdot \begin{pmatrix} 1 \\ 1 \\ 3 \\ 2 \end{pmatrix} & (2, 1, -3, 1) \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix} & (2, 1, -3, 1) \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \end{pmatrix} = \\
 &= \begin{pmatrix} 9 & 5 & 3 \\ -4 & 3 & -2 \end{pmatrix}. \tag{4.37}
 \end{aligned}$$

Method II: the matrix A multiplied by each column of B .

$$A \cdot B = \left(\begin{pmatrix} 1 & -1 & 1 & 3 \\ 2 & 1 & -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 1 & 3 \\ 2 & 1 & -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 1 & 3 \\ 2 & 1 & -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \right),$$

and one may proceed, for example, as in 6.6.3, using (6.24), thus obtaining once again (6.37).

4.7 Bilinearity and associativity of matrix multiplication

4.7.1 Proposition. *Matrix multiplication (6.6.7) is bilinear, that is, for every $A, A_1, A_2 \in \mathbb{K}^{m \times p}$, every $B, B_1, B_2 \in \mathbb{K}^{p \times n}$ and every $\lambda, \mu \in \mathbb{K}$ one has:*

$$(\lambda A_1 + \mu A_2) \cdot B = \lambda A_1 B + \mu A_2 B \tag{4.38}$$

and

$$A(\lambda B_1 + \mu B_2) = \lambda AB_1 + \mu AB_2 \tag{4.39}$$

Proof. Let $A_1, A_2 \in \mathbb{K}^{m \times p}$, $B \in \mathbb{K}^{p \times n}$ and $\lambda, \mu \in \mathbb{K}$.

4.7. BILINEARITY AND ASSOCIATIVITY OF MATRIX MULTIPLICATION 23

$$\begin{aligned}
 C_j((\lambda A_1 + \mu A_2)B) &= [(\lambda A_1 + \mu A_2)B]e_j = && \text{(formula (6.28) for the } j\text{-th} \\
 &&& \text{column of } A \cdot B) \\
 &= (\lambda A_1 + \mu A_2) \cdot (Be_j) = && \text{(by definition (6.24) of} \\
 &&& \text{matrix multiplication)} \\
 &= (\lambda A_1 + \mu A_2) \cdot C_j(B) = && \text{(formula (6.28))} \\
 &= \lambda A_1 \cdot C_j(B) + \mu A_2 \cdot C_j(B) = && \text{(linearity (6.26) of the pro-} \\
 &&& \text{duct of a matrix by a column)} \\
 &= \lambda C_j(A_1 \cdot B) + \mu C_j(A_2 \cdot B) = && \text{(definition of column and} \\
 &&& \text{of matrix multiplication)} \\
 &= C_j(\lambda A_1 \cdot B + \mu A_2 \cdot B) && \text{(second of formulas (6.16))}
 \end{aligned}$$

Since two matrices are equal if and only if they have the same columns, one concludes that

$$(\lambda A_1 + \mu A_2)B = \lambda A_1 B + \mu A_2 B,$$

whence linearity with respect to the first argument. Similarly, if $A \in \mathbb{K}^{m \times p}$ and $B_1, B_2 \in \mathbb{K}^{p \times n}$, one has:

$$\begin{aligned}
C_j(A(\lambda B_1 + \mu B_2)) &= [A(\lambda B_1 + \mu B_2)]\mathbf{e}_j = && \text{(formula (6.28) for the } j\text{-th} \\
&&& \text{column of } A \cdot B) \\
&= A((\lambda B_1 + \mu B_2)\mathbf{e}_j) = && \text{(by definition (6.24) of} \\
&&& \text{matrix multiplication)} \\
&= A \cdot C_j(\lambda B_1 + \mu B_2) = && \text{(formula (6.28))} \\
&= A(\lambda C_j(B_1) + \mu C_j(B_2)) = && \text{(second of formulas (6.16))} \\
&= \lambda AC_j(B_1) + \mu AC_j(B_2) = && \text{(linearity (6.27) of the pro-} \\
&&& \text{duct of a matrix by a column)} \\
&= \lambda A(B_1\mathbf{e}_j) + \mu A(B_2\mathbf{e}_j) = && \text{(formula (6.28) for the } j\text{-th} \\
&&& \text{column of } A \cdot B) \\
&= \lambda(A \cdot B_1)\mathbf{e}_j + \mu(A \cdot B_2)\mathbf{e}_j && \text{(Definition 6.6.7 of} \\
&&& \text{matrix multiplication)} \\
&= (\lambda(A \cdot B_1) + \mu(A \cdot B_2))\mathbf{e}_j = && (\mathbf{e}_j \text{ factored out)} \\
&= C_j(\lambda(A \cdot B_1) + \mu(A \cdot B_2)) && \text{(formula (6.28) for the} \\
&&& j\text{-th column)}
\end{aligned}$$

The equality $C_j(A(\lambda B_1 + \mu B_2)) = C_j(\lambda(A \cdot B_1) + \mu(A \cdot B_2))$, for every $1 \leq j \leq n$, proves linearity $A(\lambda B_1 + \mu B_2) = \lambda AB_1 + \mu AB_2$ of matrix multiplication with respect to the second argument. ■

4.7.2 Corollary. *The map*

$$\begin{cases} (\mathbb{K}^n)^\vee \times \mathbb{K}^n & \longrightarrow & \mathbb{K}, \\ (\alpha, \mathbf{u}) & \longmapsto & \alpha \cdot \mathbf{u}. \end{cases}$$

is bilinear, that is, linear with respect to the first and the second argument.

Proof. Obvious.

4.7.3 Proposition. *For every $(\alpha, \mathbf{u}) \in (\mathbb{K}^n)^\vee \times \mathbb{K}^n$, one has:*

$$(\alpha \mathbf{u})^T = \alpha \mathbf{u} = \mathbf{u}^T \cdot \alpha^T \quad (4.40)$$

Proof. Since a row is a $1 \times n$ matrix and a column is an $n \times 1$ matrix, the product of a row by a column is a 1×1 matrix, which of course coincides with its own transpose. This explains the first equality. The second equality consists of a direct verification. If $\alpha(i) = a_i$ and $\mathbf{u}(j) = u_j$, for $1 \leq i, j \leq n$, applying the definition

$$\begin{aligned} \alpha \cdot \mathbf{u} &= (a_1, \dots, a_n) \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = a_1 u_1 + \dots + a_n u_n = \\ &= u_1 a_1 + \dots + u_n a_n = (u_1, \dots, u_n) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \mathbf{u}^T \alpha^T. \end{aligned}$$

4.7.4 Proposition. *Matrix multiplication (6.6.7) is associative in the following sense. Let $A_1 \in \mathbb{K}^{m \times p}$, $A_2 \in \mathbb{K}^{p \times q}$ and $A_3 \in \mathbb{K}^{q \times n}$. Then:*

$$(A_1 \cdot A_2) \cdot A_3 = A_1 \cdot (A_2 \cdot A_3) \quad (4.41)$$

Proof. Let $1 \leq j \leq n$. Then:

$$\begin{aligned} C_j((A_1 \cdot A_2) \cdot A_3) &= ((A_1 \cdot A_2) \cdot A_3) \mathbf{e}_j = && \text{(definition of } j\text{-th column)} \\ &= (A_1 A_2)(A_3 \mathbf{e}_j) = && \text{(definition of the product of the} \\ &&& \text{matrices } A_1 A_2 \text{ and } A_3) \\ &= A_1(A_2 \cdot A_3 \mathbf{e}_j) = && \text{(definition of the product of } A_1 A_2 \\ &&& \text{with the column } A_3 \mathbf{e}_j) \\ &= (A_1 \cdot (A_2 \cdot A_3)) \mathbf{e}_j = && \text{(definition of the product of the matrices} \\ &&& A_1 \text{ and } A_2 \cdot A_3) \\ &= C_j(A_1 \cdot (A_2 \cdot A_3)), && \text{(definition of the } j\text{-th column} \\ &&& \text{of } A_1 \cdot (A_2 \cdot A_3)) \end{aligned}$$

and thus the equality of the columns

$$C_j((A_1 \cdot A_2) \cdot A_3) = C_j(A_1 \cdot (A_2 \cdot A_3)),$$

has been proved for every $1 \leq j \leq n$. It implies equality (6.41), which expresses the associative property. $\blacksquare$

4.7.5 Proposition. *Let $(A, B) \in \mathbb{K}^{m \times p} \times \mathbb{K}^{p \times n}$, that is, matrices for which the product makes sense. Then one has*

$$(AB)^T = B^T A^T. \quad (4.42)$$

Proof. We shall verify that for every $(i, j) \in \underline{m} \times \underline{n}$,

$$(AB)^T(i, j) = (B^T A^T)(i, j) \quad (4.43)$$

that is, that the left-hand side and the right-hand side of (6.42) have the same entries. Indeed, for every $1 \leq i \leq m$ and every $1 \leq j \leq n$:

$$\begin{aligned} (AB)^T(i, j) &= C_j(R_i((AB)^T)) = && \text{(Expression (6.6) for the entries of a matrix)} \\ &= C_j(C_i(AB)^T) = && \text{(Formula (6.12) for the row of a transpose matrix)} \\ &= [R_j(C_i(AB))]^T = && \text{(Formula (6.13) for the column of a transpose matrix)} \\ &= (AB(j, i))^T = && \text{(Definition of } j\text{-th row and } i\text{-th column)} \\ &= (R_j(A)C_i(B))^T = && \text{(formula (6.33) for the product of matrices)} \\ &= C_i(B)^T R_j(A)^T = && \text{(equalities (6.40) on the transpose of the product row by column)} \\ &= R_i(B^T)C_j(A^T) = && \text{(again formulas (6.12) and (6.13))} \\ &= (B^T \cdot A^T)(i, j). && \text{(equality (6.33) applied to the product of matrices } B^T A^T) \end{aligned}$$

Equality (6.43) is thus proved for every $(i, j) \in \underline{m} \times \underline{n}$, and therefore (6.42) is also proved. ■

The conjugate of the product AB of two matrices “behaves better”:

4.7.6 Proposition. For every pair $(A, B) \in \mathbb{K}^{m \times p} \times \mathbb{K}^{p \times n}$ one has:

$$\overline{AB} = \overline{A} \cdot \overline{B} \quad (4.44)$$

In fact

$$\overline{AB}(i, j) = \overline{AB(i, j)} = \overline{R_i(A)C_j(B)}.$$

Expanding the product of the i -th row with the j -th column, one gets

$$\overline{R_i(A)C_j(B)} = \overline{\sum_{h=1}^p R_i(A)(h) \cdot C_j(B)(h)} = \overline{\sum_{h=1}^p A(i, h) \cdot B(h, j)}.$$

The latter sum, by the properties of conjugation, is equal to

$$\sum_{h=1}^p \overline{A(i, h)} \cdot \overline{B(h, j)} = \sum_{h=1}^p \overline{A}(i, h) \cdot \overline{B}(h, j) = R_i(\overline{A}) \cdot C_j(\overline{B})$$

which is precisely the entry $\overline{A} \cdot \overline{B}(i, j)$ of the product $\overline{A} \cdot \overline{B}$ of the conjugate of A by the conjugate of B . $\blacksquare$

4.8 Square matrices

4.8.1 A matrix $A \in \mathbb{K}^{n \times n}$, that is, one having as many rows as columns, is said to be *square*. The space $\mathbb{K}^{1 \times 1}$ of 1×1 matrices can be identified with $\mathbb{K}$ through the bijection $(a) \mapsto a$ that *forgets* the parentheses. Moreover matrix multiplication induces an internal binary operation:

$$\mathbb{K}^{n \times n} \times \mathbb{K}^{n \times n} \longrightarrow \mathbb{K}^{n \times n}$$

which associates with each pair (A, B) of $n \times n$ matrices the product matrix AB .

4.8.2 Matrix multiplication is commutative only for $n = 1$ (1×1 matrices). To prove that for $n \geq 2$ the product is not commutative, it is sufficient to exhibit for every n a pair of matrices that do not commute. For example, for every $n \geq 2$ consider the matrices

$$A_n = (\mathbf{e}_2, \mathbf{e}_1, \mathbf{u}_3, \dots, \mathbf{u}_n) \quad \text{and} \quad B_n = (\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1, \mathbf{u}_3, \dots, \mathbf{u}_n).$$

We claim that $A_n \cdot B_n \neq B_n \cdot A_n$. To show this it is sufficient to show that the matrices $A_n B_n$ and $B_n A_n$ do not have the same columns. Now:

$$C_1(A_n B_n) = A_n \cdot C_1(B_n) = A_n(\mathbf{e}_1 + \mathbf{e}_2) = A_n \mathbf{e}_1 + A_n \mathbf{e}_2 = C_1(A_n) + C_2(A_n) = \mathbf{e}_1 + \mathbf{e}_2,$$

whereas

$$C_1(BA) = B \cdot C_1(A) = B\mathbf{e}_2 = C_2(B) = \mathbf{e}_1.$$

Since $C_1(A_n B_n) \neq C_1(B_n A_n)$, it follows that $A_n B_n \neq B_n A_n$.

4.8.3 Esempio. For the skeptical reader, or for the one who likes to do (more) explicit calculations, for $n = 2$ one has

$$A_2 = (\mathbf{e}_2, \mathbf{e}_1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad e \quad B_2 = (\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

and therefore:

$$A_2 B_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = B_2 A_2.$$

4.8.4 The product in $\mathbb{K}^{n \times n}$ is of course associative, because matrix multiplication is associative in general. Moreover the identity matrix $\mathbf{1}_n$ defined by $\mathbf{1}_n(i, j) = \delta_{ij}$, or, more explicitly:

$$\mathbf{1}_n = \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_n \end{pmatrix} = (\mathbf{e}_1, \dots, \mathbf{e}_n),$$

is the neutral element for the product (of matrices). Indeed, for every $A \in \mathbb{K}^{n \times n}$:

$$A \cdot \mathbf{1}_n = A(\mathbf{e}_1, \dots, \mathbf{e}_n) = (A\mathbf{e}_1, \dots, A\mathbf{e}_n) = (C_1(A), \dots, C_n(A)) = A$$

and

$$\mathbf{1}_n \cdot A = \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_n \end{pmatrix} \cdot A = \begin{pmatrix} \mathbf{e}_1 A \\ \vdots \\ \mathbf{e}_n A \end{pmatrix} = \begin{pmatrix} R_1(A) \\ \vdots \\ R_n(A) \end{pmatrix} = A.$$

The zero matrix defined by $\mathbf{0}(i, j) = 0$ is instead an annihilator for the product, that is

$$\mathbf{0}_n \cdot A = A \cdot \mathbf{0}_n = \mathbf{0}_n$$

for every $A \in \mathbb{K}^{n \times n}$. In this case the trivial verification is left to the reader.

4.8.5 The set $\mathbb{K}^{n \times n}$ is therefore a vector space on which a product structure is defined (the product of square matrices), which is bilinear, in the sense already clarified in Definition 6.7.1. One then says that $\mathbb{K}^{n \times n}$ is an *algebra*, called the *$\mathbb{K}$ -algebra of square $n \times n$ matrices*. Indeed $\mathbb{K}^{n \times n}$ is an abelian group with respect to addition, matrix product is associative, with neutral element. Moreover the distributive properties of the product with respect to addition hold:

$$(A_1 + A_2) \cdot B = A_1 B + A_2 B \quad \text{and} \quad A(B_1 + B_2) = AB_1 + AB_2,$$

equalities that need not be proved. They are indeed a particular case of the bilinearity (6.7.1) of matrix multiplication in general. Hence $\mathbb{K}^{n \times n}$ is a ring, but not a commutative one, since $AB \neq BA$ in general (as was seen in the counterexample 6.8.3).

4.8.6 Esempio. The $\mathbb{R}$ -algebra $\mathbb{R}^{2 \times 2}$ of real 2×2 square matrices has a commutative subalgebra which, for reasons that will be clear in chapter ??, we shall dare to call *Cauchy-Riemann*. It is the subalgebra $\mathcal{CR}$ of all square matrices $A \in \mathbb{R}^{2 \times 2}$ such that

$$\begin{cases} A(1,1) - A(2,2) = 0 \\ A(1,2) + A(2,1) = 0 \end{cases}$$

A generic such matrix is of the form

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a\mathbf{1}_{2 \times 2} + b\mathbb{I}, \quad a, b \in \mathbb{R}.$$

where

$$\mathbb{I} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \Bigg| \quad \mathbb{I}^2 = -\mathbf{1}_{2 \times 2}$$

Since

$$\lambda(a\mathbf{1}_{2 \times 2} + b\mathbb{I}) + \mu(a'\mathbf{1}_{2 \times 2} + b'\mathbb{I}) = (\lambda a + \mu a')\mathbf{1}_{2 \times 2} + (\lambda b + \mu b')\mathbb{I},$$

$\mathcal{CR}$ is a vector subspace of $\mathbb{R}^{2 \times 2}$. Moreover the product is commutative:

$$(a\mathbf{1}_{2 \times 2} + b\mathbb{I}) \cdot (a'\mathbf{1}_{2 \times 2} + b'\mathbb{I}) = (a'\mathbf{1}_{2 \times 2} + b'\mathbb{I}) \cdot (a\mathbf{1}_{2 \times 2} + b\mathbb{I}) \quad (4.45)$$

as the female and male reader will easily verify, by expanding separately the left-hand side and the right-hand side of (6.45). What does the result resemble?

4.8.7 On the vector space $\mathbb{K}^{n \times n}$ one may define a further noncommutative algebra structure. For every pair $(A, B) \in \mathbb{K}^{n \times n}$ define the *commutator* of A and B through the equality

$$[A, B] = AB - BA. \quad (4.46)$$

The reason for the terminology is that two square matrices $A, B \in \mathbb{K}^{n \times n}$ commute (that is $AB = BA$) if and only if $[A, B] = 0$. The verification that the map

$$\begin{cases} [\cdot, \cdot] : \mathbb{K}^{n \times n} \times \mathbb{K}^{n \times n} \longrightarrow \mathbb{K}^{n \times n} \\ (A, B) \longmapsto [A, B] \end{cases}$$

is bilinear and antisymmetric is left to the reader, that is, for every choice of $A, A_1, A_2, B, B_1, B_2 \in \mathbb{K}^{n \times n}$ and $\lambda, \mu \in \mathbb{K}$ one has :

$$[\lambda A_1 + \mu A_2, B] = \lambda[A_1, B] + \mu[A_2, B] \quad (4.47)$$

$$[A, \lambda B_1 + \mu B_2] = \lambda[A, B_1] + \mu[A, B_2] \quad (4.48)$$

which express linearity with respect to the first and second argument, respectively, and

$$[A, B] = -[B, A], \quad (4.49)$$

which expresses antisymmetry. The algebra $\mathbb{K}^{n \times n}$ with respect to the commutator $[\cdot, \cdot]$ is not associative. Indeed for a general choice of matrices $A, B, C \in \mathbb{K}^{n \times n}$ one has:

$$[[A, B], C] \neq [A, [B, C]].$$

In fact $(\mathbb{K}^{n \times n}, [\cdot, \cdot])$ is one of the simplest examples, besides the already seen one of $\mathbb{R}^3$ with respect to the cross product, of a *Lie algebra*, since it satisfies the famous *Jacobi identity*

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0 \quad (4.50)$$

whose simple verification is left once again to the reader.

4.9 The linear group of invertible matrices

As has been seen, the set $\mathbb{K}^{n \times n}$, with respect to addition and multiplication of square matrices, is a $\mathbb{K}$ -*algebra* (noncommutative). It is natural to ask whether it may also have the structure of a *division ring*², that is, whether every nonzero square matrix is invertible. One will immediately see that this is not so.

4.9.1 Definition. A square matrix $A \in \mathbb{K}^{n \times n}$ is said to be invertible if and only if there exists $B \in \mathbb{K}^{n \times n}$ such that

$$AB = BA = \mathbf{1}_n.$$

In this case B is said to be an *inverse* of A . Moreover A is the inverse of the matrix B , since similarly $BA = AB = \mathbf{1}_n$. An invertible matrix A admits a unique inverse.

²A *field* is a commutative *division ring*, that is, with commutative multiplication.

Indeed, if B' were another inverse of A , one would have

$$BAB' = (BA)B' = \mathbf{1}_n B' = B'$$

||

$$BAB' = B(AB') = B\mathbf{1}_n = B$$

whence $B' = B$ (because the product of B , A , and B' has a unique result).

4.9.2 Notation. In the case where the matrix A is invertible, its unique inverse will be denoted from now on by the symbol A^{-1} . If B is the inverse of A , then $B^{-1} = A$, since:

$$BA = AB = \mathbf{1}_n,$$

having used the fact that A is the inverse of B .

4.9.3 The zero square matrix $\mathbf{0}_n$ is obviously non-invertible, being an annihilator of the product. However it is not the only one. For example the matrix

$$A := (\mathbf{e}_1, \mathbf{0}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

is nonzero but not invertible. Indeed, if it had an inverse, say B , one should have

$$BA = \mathbf{1}_2 = (\mathbf{e}_1, \mathbf{e}_2),$$

that is, in particular, $C_2(BA) = \mathbf{e}_2$. This is impossible, since:

$$C_2(BA) = BC_2(A) = B \cdot \mathbf{0} = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

4.9.4 Exercise Let $\mathbf{u} \in \mathbb{K}^2$ be a nonzero vector and $\lambda \neq 0$ a scalar. Prove that the matrix

$$A := (\mathbf{u}, \lambda \mathbf{u})$$

is not invertible.

4.9.5 Proposition. *The product of two invertible matrices is invertible.*

Proof. Let A and B be invertible matrices. One will show that the inverse of the matrix AB is the matrix $B^{-1}A^{-1}$, that is:

$$(AB)^{-1} = B^{-1}A^{-1} \tag{4.51}$$

Indeed

$$\begin{aligned}
AB(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = && \text{(Associative property (6.41) of matrix multiplication)} \\
&= A\mathbf{1}_n A^{-1} = && \text{(definition of inverse matrix)} \\
&= (A\mathbf{1}_n)A^{-1} = && \text{(Associative property (6.41) of matrix multiplication)} \\
&= AA^{-1} = \mathbf{1}_n, && \text{(definition of inverse matrix).}
\end{aligned}$$

Similarly, and this time omitting the explanation of the steps, entirely analogous after all:

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}\mathbf{1}_n B = B^{-1}(\mathbf{1}_n B) = B^{-1}B = \mathbf{1}_n. \quad \blacksquare$$

4.9.6 Later on it will be seen, in general, that a square matrix $A \in \mathbb{K}^{n \times n}$ is invertible if and only if its columns form a basis of the space $\mathbb{K}^n$, if and only if its rows form a basis of $(\mathbb{K}^n)^\vee$.

4.9.7 Proposition. *Let A be an invertible matrix. Then one has:*

$$(A^T)^{-1} = (A^{-1})^T. \quad (4.52)$$

In other words, the *inverse of the transpose* coincides with the *transpose of the inverse*.

Proof. Since $AA^{-1} = \mathbf{1}_n$, one has

$$(AA^{-1})^T = (\mathbf{1}_n)^T = \mathbf{1}_n.$$

Applying (6.42) one has:

$$(A^{-1})^T A^T = \mathbf{1}$$

whence $(A^T)^{-1} = (A^{-1})^T$. ■

Similarly:

4.9.8 Proposition. *If $A \in \mathbb{K}^{n \times n}$ is invertible, then the conjugate $\overline{A}$ is invertible and its inverse is the conjugate of the inverse:*

$$(\overline{A})^{-1} = \overline{A^{-1}}. \quad (4.53)$$

Proof. Indeed:

$$\overline{A^{-1}} \cdot \overline{A} = \overline{A^{-1}A} = \overline{\mathbf{1}_n} = \mathbf{1}_n,$$

and

$$\overline{A} \cdot \overline{A^{-1}} = \overline{A \cdot A^{-1}} = \overline{\mathbf{1}_n} = \mathbf{1}_n.$$

whence (6.53).

4.9.9 Proposition/Definition. *The set*

$$Gl_n(\mathbb{K}) := \{A \in \mathbb{K}^{n \times n} \mid A \text{ is invertible}\}$$

of all invertible $n \times n$ square matrices is a group with respect to matrix multiplication, called the linear group of $\mathbb{K}^n$.

Proof. Matrix multiplication is an internal binary operation. Indeed, according to Proposition 6.9.9, the product of two invertible matrices is invertible. Moreover the identity matrix, neutral element of the product of square matrices, is invertible (since it coincides with its inverse) and therefore $Gl_n(\mathbb{K})$ contains the neutral element of the product. By construction, every $A \in Gl_n(\mathbb{K})$ has an inverse. Hence $Gl_n(\mathbb{K})$ is a group, noncommutative in view of what was observed in 6.8.2. ■

4.10 Symmetric and antisymmetric matrices

4.10.1 A matrix $A \in \mathbb{K}^{n \times n}$ is said to be symmetric if $A^T = A$, that is, if and only if the matrix A^T obtained from A by exchanging rows and columns coincides with A .

4.10.2 Esempio. Consider the matrices A and B that follow:

$$A := \begin{pmatrix} 1 & 2 & 3 & 1 \\ 2 & -3 & 4 & 1 \\ 3 & 4 & -5 & 2 \\ 1 & 1 & 2 & 5 \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} 1 & 2 & 3 & 1 \\ 2 & 3 & 1 & 4 \\ 3 & 5 & -5 & 2 \\ 1 & 1 & 9 & 1 \end{pmatrix}$$

Then A is symmetric (it does not change when rows and columns are exchanged) and B is not symmetric.

4.10.3 If A, B are symmetric matrices, it is not necessarily true that AB is a symmetric matrix. Indeed

$$(AB)^T = B^T A^T = BA,$$

where the first equality follows from (6.42) and the second from the fact that, by hypothesis, $A = A^T$ and $B = B^T$. The product AB of two symmetric matrices is therefore symmetric if and only if $[A, B] = 0$ (Cf. 6.46), that is, if and only if the matrices A and B commute. However the inverse of an invertible symmetric matrix is a symmetric matrix. Indeed, if $A = A^T$:

$$(A^{-1})^T = (A^T)^{-1} = A^{-1}, \quad (4.54)$$

where the first equality comes from Proposition 6.9.7, while the second comes from the hypothesis $A^T = A$.

4.10.4 Proposition. *A matrix $A \in \mathbb{K}^{n \times n}$ is said to be antisymmetric if and only if*

$$A^T = -A. \quad (4.55)$$

In particular, observing that transposition does not “move” the elements of the main diagonal, one has, for every $1 \leq i \leq n$:

$$A(i, i) = A^T(i, i) = -A(i, i),$$

whence one deduces that $A(i, i) = 0$ for every $1 \leq i \leq n$. That is: all elements of the main diagonal of an antisymmetric matrix are zero.

4.10.5 Esempio. The matrix

$$\Omega := \begin{pmatrix} 0 & \omega_1 & -\omega_2 \\ -\omega_1 & 0 & \omega_3 \\ \omega_2 & -\omega_3 & 0 \end{pmatrix}$$

is antisymmetric. Let the reader try to write down a 4×4 antisymmetric matrix.

4.10.6 Proposition. *Let $\mathcal{S}^{n \times n}$ and $\mathcal{A}^{n \times n}$ be, respectively, the subsets of symmetric and antisymmetric square matrices. Then:*

- a) $\mathcal{S}^{n \times n}$ and $\mathcal{A}^{n \times n}$ are vector subspaces of $\mathbb{K}^{n \times n}$;
- b) The space of $n \times n$ square matrices with entries in $\mathbb{K}$ is the direct sum of the subspaces $\mathcal{S}$ and $\mathcal{A}$:

$$\mathbb{K}^{n \times n} = \mathcal{S}^{n \times n} \oplus \mathcal{A}^{n \times n}. \quad (4.56)$$

Proof. The proof of a) is easy and is based on the properties of transposition. Indeed, if A, B are symmetric and $\lambda, \mu \in \mathbb{K}$ are arbitrary,

$$(\lambda A + \mu B)^T = \lambda A^T + \mu B^T = \lambda A + \mu B,$$

that is, an arbitrary linear combination of two symmetric matrices is a symmetric matrix. Similarly, if $A, B \in \mathcal{A}$:

$$(\lambda A + \mu B)^T = \lambda A^T + \mu B^T = \lambda(-A) + \mu(-B) = -(\lambda A + \mu B),$$

that is, an arbitrary linear combination of two antisymmetric matrices is an antisymmetric matrix. As for b), observe that if $M \in \mathbb{K}^{n \times n}$, then

$$M_{\mathcal{S}} = \frac{1}{2}(M + M^T) \quad \text{and} \quad M_{\mathcal{A}} = \frac{1}{2}(M - M^T)$$

are two matrices, respectively symmetric and antisymmetric. Moreover

$$M = M_{\mathcal{S}} + M_{\mathcal{A}}.$$

The last equality proves that every matrix can be written in at least one way as the sum of a symmetric matrix and an antisymmetric one. Hence $\mathbb{K}^{n \times n}$ is the sum of the subspaces $\mathcal{S}$ and $\mathcal{A}$. To prove that the sum is direct it is enough to prove that $\mathcal{S} \cap \mathcal{A} = \{\mathbf{0}_{n \times n}\}$, that is, that the only matrix both symmetric and antisymmetric is the zero matrix. But if $M \in \mathcal{S} \cap \mathcal{A}$, then:

$$-M = M^T = M,$$

where the first equality comes from the fact that M is antisymmetric and the second from the fact that it is symmetric. Therefore $M = -M$, that is $M = \mathbf{0}_{n \times n}$ ³ ■

4.10.7 Definition. A matrix $A \in \mathbb{K}^{n \times n}$ is said to be Hermitian if and only if

$$A^T = \overline{A}$$

It is said to be anti-Hermitian if and only if

$$A^T = -\overline{A} \tag{4.57}$$

Clearly real symmetric and antisymmetric matrices are a particular case of Hermitian and anti-Hermitian matrices. Indeed, if the matrices have real coefficients, conjugation leaves the matrix unchanged. It is not true, however, that a Hermitian matrix is symmetric, as shown by the matrix

$$H := \begin{pmatrix} 3 & 2 - 3i \\ 2 + 3i & 5 \end{pmatrix}$$

which is Hermitian but not symmetric. We conclude this survey of matrices with the following definitions.

³Again because $2 \neq 0$ in the fields we are considering.

4.10.8 Definition. A matrix $A \in \mathbb{K}^{n \times n}$ is said to be unitary if and only if

$$A\overline{A}^T = \overline{A}^T A = \mathbf{1}_n. \quad (4.58)$$

Needless to say, a unitary matrix is invertible, having as inverse the conjugate of the transpose (which is the same thing as the transpose of the conjugate). If $\mathbb{K} = \mathbb{R}$, equation (6.58) rewrites more simply as:

$$A \cdot A^T = A^T \cdot A = \mathbf{1}_n$$

In that case, A is called a *real orthogonal matrix*.

4.11 Rank, reduced matrices, and reduction

We shall begin immediately with a fundamental:

4.11.1 Definition. The row space (or space of rows) of A is

$$[\mathbf{R}_1(A), \dots, \mathbf{R}_m(A)],$$

that is, the vector subspace of $(\mathbb{K}^n)^\vee$ generated by the rows of A . The column space (or space of columns) of A is

$$[\mathbf{C}_1(A), \dots, \mathbf{C}_n(A)]$$

that is, the vector subspace of $\mathbb{K}^m$ generated by the columns of A .

4.11.2 Esempio. Let

$$A := \begin{pmatrix} 1 & 2 & 1 & 5 \\ 1 & 3 & -1 & 2 \\ 5 & 13 & -1 & 16 \end{pmatrix}.$$

Then the row space is:

$$[\mathbf{R}_1(A), \mathbf{R}_2(A), \mathbf{R}_3(A)] = [(1, 2, 1, 5), (1, 3, -1, 2), (5, 13, -1, 16)].$$

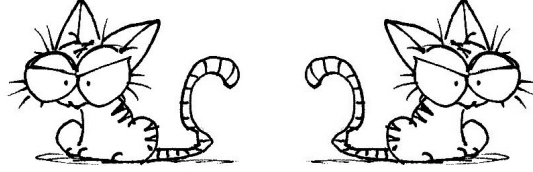
Since, however, the third row is a linear combination of the first two:

$$\mathbf{R}_3(A) = 2\mathbf{R}_1(A) + 3\mathbf{R}_2(A),$$

it follows that the third row is irrelevant for the purpose of describing the row space. Therefore $[\mathbf{R}_1(A), \mathbf{R}_2(A)]$ is the row space of A . Clearly the same space is generated by any two of the rows of A . The column space of A is instead:

$$[\mathbf{C}_1(A), \mathbf{C}_2(A)] = \left[\begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 13 \end{pmatrix} \right]$$

since all the other columns, as the reader will take care to verify, are linear combinations of these first two. Clearly the row space is a subspace of $(\mathbb{R}^4)^\vee$, whereas the column space is a subspace of $\mathbb{R}^3$. Note that although they are different vector spaces, they have the same dimension (equal to two).



4.11.3 Proposition. *Let $A \in \mathbb{K}^{m \times n}$. Then*

$$\dim_{\mathbb{K}}[R_1(A), \dots, R_m(A)] = \dim_{\mathbb{K}}[C_1(A), \dots, C_n(A)],$$

that is: the dimension of the row space of A coincides with the dimension of the column space of A .

Proof. The proof is not long, but the matrices take up a lot of space. Let $r = \dim_{\mathbb{K}}[R_1(A), \dots, R_m(A)]$ and $r' = \dim_{\mathbb{K}}[C_1(A), \dots, C_n(A)]$. One will show that $r \leq r'$ and $r' \leq r$, whence the equality of r and r' . By hypothesis, there exists at least one r -tuple of rows of A that is linearly independent. It is no restriction to suppose that the first r rows are linearly independent, up to a reordering of the rows themselves. Since they are, by definition, a basis of the row space, one has:

$$\begin{pmatrix} R_1(A) \\ R_2(A) \\ \vdots \\ R_r(A) \\ R_{r+1}(A) \\ \vdots \\ R_m(A) \end{pmatrix} = \begin{pmatrix} R_1(A) \\ R_2(A) \\ \vdots \\ R_r(A) \\ a_{r+1,1}R_1(A) + a_{r+1,2}R_2(A) + \dots + a_{r+1,r}R_r(A) \\ \vdots \\ a_{n,1}R_1(A) + a_{n,2}R_2(A) + \dots + a_{n,r}R_r(A) \end{pmatrix}$$

where the last $m - r$ rows have been expressed as linear combinations of the first r . Therefore, the j -th column of A , $1 \leq j \leq n$, is given by:

$$\begin{aligned} C_j(A) = A\mathbf{e}_j &= \begin{pmatrix} R_1(A) \\ R_2(A) \\ \vdots \\ R_r(A) \\ R_{r+1}(A) \\ \vdots \\ R_m(A) \end{pmatrix} \cdot \mathbf{e}_j = \begin{pmatrix} R_1(A)\mathbf{e}_j \\ R_2(A)\mathbf{e}_j \\ \vdots \\ R_r(A)\mathbf{e}_j \\ a_{r+1,1}R_1(A)\mathbf{e}_j + \dots + a_{r+1,r}R_r(A)\mathbf{e}_j \\ \vdots \\ a_{n,1}R_1(A)\mathbf{e}_j + \dots + a_{n,r}R_r(A)\mathbf{e}_j \end{pmatrix} = \\ &= R_1(A)\mathbf{e}_j \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ a_{r+1,1} \\ \vdots \\ a_{n,1} \end{pmatrix} + R_2(A)\mathbf{e}_j \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ a_{r+1,2} \\ \vdots \\ a_{n,2} \end{pmatrix} + \dots + R_r(A)\mathbf{e}_j \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ a_{r+1,r} \\ \vdots \\ a_{n,r} \end{pmatrix}. \end{aligned}$$

Thus it has been proved that every column of A belongs to the vector subspace of $\mathbb{K}^n$ generated by the r columns:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ a_{r+1,1} \\ \vdots \\ a_{n,1} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ a_{r+1,2} \\ \vdots \\ a_{n,2} \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ a_{r+1,r} \\ \vdots \\ a_{n,r} \end{pmatrix},$$

and therefore the dimension r' of the column space is at most r , that is $r' \leq r$. Now observe that the row space of the matrix A^T has dimension r' (because the rows are the transposed columns). Moreover, the dimension of the column space of A^T is r (because the columns of A^T are the rows of A). Then, by the first part of the proof, one also has the equality $r \leq r'$, whence $r = r'$, as was to be proved. ■

4.11.4 Definition. The rank of $A \in \mathbb{K}^{m \times n}$ is the dimension of the row space or, equivalently, of the column space. If r is the rank of a matrix, we shall write:

$$\text{rk}_{\mathbb{K}}(A) = r$$

and therefore

$$\dim_{\mathbb{K}}[C_1(A), \dots, C_n(A)] = \dim_{\mathbb{K}}[R_1(A), \dots, R_m(A)].$$

4.11.5 Esempio. The subscript “ $\mathbb{K}$ ” appended to the symbol rk for rank may be omitted only when the field is clear from the context. Indeed, let A be the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ i & i & i \end{pmatrix}$$

then $\text{rk}_{\mathbb{C}} A = 1$, since $R_2(A) = iR_1(A)$ and therefore $R_1(A)$ and $R_2(A)$ are linearly dependent (over the complex numbers). However over the reals the rows $R_1(A)$ and $R_2(A)$ are linearly independent, and therefore $\text{rk}_{\mathbb{R}}(A) = 2$.

4.11.6 The goal is now to set up an algorithm that allows one to determine bases of the row space and of the column space by means of a suitable succession of *elementary transformations*. Two matrices $A, B \in \mathbb{K}^{m \times n}$ are said to be row-equivalent if and only if the subspace of $(\mathbb{K}^n)^\vee$ generated by the rows of A coincides with the one generated by the rows of B . More prosaically, A and B are row-equivalent if and only if they have the same row space. Note that the row spaces may be equal without the matrices themselves being equal: this only means that the rows of one and of the other form distinct generating sets of the same vector subspace. Similarly two matrices A and B of $\mathbb{K}^{m \times n}$ are said to be column-equivalent if they have the same column space, that is, if and only if their transposes are row-equivalent. There are two types of elementary transformations: the so-called *row* ones and the so-called *column* ones. The former transform a matrix into another one row-equivalent to it, and the latter transform the matrix into another one column-equivalent to it. It should be observed that two matrices row-equivalent to each other might not have the same column space.

4.11.7 Esempio. Consider the two matrices

$$A := \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 & 1 \\ 3 & 3 & 3 \end{pmatrix}.$$

The two matrices A and B , both of rank 1, have the same row space, each being generated by the row $(1, 1, 1)$. However the column space of A is generated by the column

$$\mathbf{u}_A := \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

whereas that of the matrix B is generated by the column

$$\mathbf{u}_B := \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Clearly $\mathbf{u}_A \neq \mathbf{u}_B$ and, thus, one sees that being row-equivalent does not mean being column-equivalent.

From every matrix $A \in \mathbb{K}^{m \times n}$ one may obtain a row-equivalent (or column-equivalent) matrix in *reduced* form.

4.11.8 Definition. *A matrix $A \in \mathbb{K}^{m \times n}$ is said to be row-reduced if and only if every nonzero row has at least one nonzero element below which, in the same column, there are only zeros or no further entries.*

4.11.9 Esempio. The matrix

$$\begin{pmatrix} 1 & 2 & -1 & 5 & 9 \\ 2 & 0 & -1 & 0 & 9 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix}$$

is row-reduced. Indeed the first row (which is nonzero) has exactly two nonzero elements (the 2 and the 5) below which there are only zeros. The second row (which is nonzero) has the element 9 below which there are only zeros. The third row (which is nonzero) has the entry 1 below which there is a zero. The fourth row has one nonzero element, namely 1. Below that one there are no zeros, because the matrix has no more rows (below that one there are no further entries). The matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & -1 & 5 & 9 \\ 2 & 0 & -1 & 0 & 9 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix}$$

is also row-reduced. It is true that the first row has no element below which there are only zeros. On the other hand the first row is zero, and therefore it satisfies the definition trivially (below all the nonzero elements⁴ of the first row there are only zeros). The matrix

$$\begin{pmatrix} 1 & 2 & -1 & 5 & 9 \\ 2 & 0 & -1 & 1 & 9 \\ 1 & 3 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix}$$

is not reduced. The first row has no element below which there are only zeros.

⁴That is, none!

Every element in each row below which there are only zeros (or no other entry) is called a *special element*. In the matrix of Example 6.11.9, the special elements are those highlighted by the little box in the following display:

$$\begin{pmatrix} 1 & \boxed{2} & -1 & \boxed{5} & 9 \\ 2 & 0 & -1 & 0 & \boxed{9} \\ \boxed{1} & 0 & -1 & 0 & 0 \\ 0 & 0 & \boxed{-1} & 0 & 0 \end{pmatrix}$$

4.11.10 Definition. A matrix is said to be column-reduced if and only if its transpose is row-reduced.

Equivalently one might say that a matrix $A \in \mathbb{K}^{m \times n}$ is column-reduced if and only if every nonzero column has a nonzero element which has only zeros to its right. More formally, A is column-reduced if and only if for every $1 \leq j \leq n$ such that $C_j(A) \neq 0$, there exists $1 \leq i_j \leq m$ such that $A(i_j, j) \neq 0$ and $A(i, j) = 0$ for every $i_j < i \leq n$.

4.11.11 Esempio. The matrix

$$B = \begin{pmatrix} 3 & 1 & \mathbf{2} & 0 \\ \mathbf{1} & 0 & 0 & 0 \\ 0 & \mathbf{2} & 0 & 0 \end{pmatrix}$$

is column-reduced. The special elements are indicated in boldface.

4.11.12 Proposition. The rank of a matrix A that is row-reduced (respectively column-reduced) coincides with the number of nonzero rows.

Proof. It is enough to prove the proposition for column-reduced matrices, since if a matrix is row-reduced its transpose is column-reduced. Suppose that the matrix A , column-reduced, has r nonzero columns. It is no restriction to assume that these are the first r , $C_1(A), \dots, C_r(A)$. It will suffice to prove that they are linearly independent. To this end consider a zero linear combination

$$\lambda_1 C_1(A) + \dots + \lambda_r C_r(A) = \mathbf{0}_m. \quad (4.59)$$

Let $i_1, i_2, \dots, i_r$ be the special elements of the first, second, $\dots$, r -th column. Column reduction means that $C_j(A)(i_h) = 0$ for every $h < j$. Thus (6.60) implies

$$\lambda_1 C_1(A)(i_1) + \lambda_2 C_2(A)(i_1) + \dots + \lambda_r C_r(A)(i_1) = \mathbf{0}_m \quad (4.60)$$

whence, by the reduction hypothesis, $C_j(A)(i_1) = 0$ for all $j \geq 2$. Therefore (6.60) is equivalent to

$$\lambda_1 C_1(A)(i_1) = 0$$

whence $\lambda_1 = 0$, since by hypothesis $C_1(A)(i_1) \neq 0$. Suppose one has proved that $\lambda_i = 0$ for all $i \leq h-1 \leq r-1$. Then

$$\lambda_h C_h(A) + \dots + \lambda_r C_r(A) = \mathbf{0}_m$$

Considering the i_h -th component of the previous sum one obtains:

$$0 = \lambda_h C_h(A)(i_h) + \dots + \lambda_r C_r(A)(i_h) = \lambda_h C_h(A)(i_h)$$

whence, since $C_h(A)(i_h) \neq 0$, one obtains $\lambda_h = 0$. By induction one concludes that all coefficients of the zero linear combination (6.60) are zero. Therefore the r columns are linearly independent, and the rank of the reduced matrix A coincides precisely with the number of nonzero columns. The same result holds for rows by observing that if A is row-reduced, its transpose is column-reduced. ■



-- Columns? --

Proposition 6.11.12 therefore suggests a method for computing the rank of a matrix with many rows or many columns, namely to compute it on a row-reduced or column-reduced equivalent matrix. Fortunately there exists such a procedure, called reduction, which we now proceed to describe by means of elementary matrices.

4.12 Elementary matrices and reduction

4.12.1 Consider the set $\mathcal{E}$ of all $m \times m$ square matrices of the form:

$$\mathcal{E}_{1,i}(\lambda) = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \lambda \epsilon_i \\ \vdots \\ \epsilon_m \end{pmatrix} \quad (\lambda \neq 0), \quad \mathcal{E}_{2,i,j}(\lambda) = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_i + \lambda \epsilon_j \\ \vdots \\ \epsilon_m \end{pmatrix}, \quad \mathcal{E}_{3,i,j} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_j \\ \vdots \\ \epsilon_i \\ \vdots \\ \epsilon_m \end{pmatrix} \quad (i < j).$$

They are called *elementary row matrices*. The matrix $\mathcal{E}_{1,i}(\lambda)$ is obtained from the identity $\mathbf{1}_m \in \mathbb{K}^{m \times m}$ by multiplying its i -th row by $0 \neq \lambda \in \mathbb{K}$; the matrix $\mathcal{E}_{2,i,j}(\lambda)$ is obtained from the identity $\mathbf{1}_m$ by adding to the i -th row the j -th row multiplied by λ ; finally the matrix $\mathcal{E}_{3,i,j}$ is obtained from the identity $\mathbf{1}_m$ by exchanging rows i and j .

4.12.2 Esempio. Here is an illustration of three elementary matrices in $\mathbb{K}^{4 \times 4}$:

$$\mathcal{E}_{1,3}(5) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{E}_{2;2,3}(-2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{E}_{3;3,4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Similarly, there are *elementary column matrices*:

$$\begin{aligned} \mathcal{F}_{1,i}(\lambda) &= (\mathbf{e}_1, \dots, \lambda \mathbf{e}_i, \dots, \mathbf{e}_n) \\ \mathcal{F}_{2;i,j}(\lambda) &= (\mathbf{e}_1, \dots, \mathbf{e}_i + \lambda \mathbf{e}_j, \dots, \mathbf{e}_n) \\ \mathcal{F}_{3;i,j} &= (\mathbf{e}_1, \dots, \mathbf{e}_j, \dots, \mathbf{e}_i, \dots, \mathbf{e}_n) \quad (i < j). \end{aligned}$$

The matrix $\mathcal{F}_{1,i}(\lambda)$ is obtained from the identity $\mathbf{1}_n$ by multiplying its i -th column by $0 \neq \lambda \in \mathbb{K}$; the matrix $\mathcal{F}_{2;i,j}(\lambda)$ is obtained from the identity $\mathbf{1}_n$ by adding to the i -th column λ times its j -th column; finally the matrix $\mathcal{F}_{3;i,j}$ is obtained from the identity $\mathbf{1}_n$ by exchanging columns i and j .

4.12.3 Esempio. Here are three examples of elementary column matrices in $\mathbb{K}^{4 \times 4}$:

$$\mathcal{F}_{1,3}(5) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{F}_{2;2,3}(-2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{F}_{3;1,4} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

4.12.4 Now let $A \in \mathbb{K}^{m \times n}$ and consider the effect of multiplying it on the left by a matrix in $\mathcal{E}$:

$$\mathcal{E}_{1,i}(\lambda) \cdot A = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \lambda \epsilon_i \\ \vdots \\ \epsilon_m \end{pmatrix} \cdot A = \begin{pmatrix} \epsilon_1 \cdot A \\ \vdots \\ \lambda \epsilon_i \cdot A \\ \vdots \\ \epsilon_m \cdot A \end{pmatrix} = \begin{pmatrix} R_1(A) \\ \vdots \\ \lambda R_i(A) \\ \vdots \\ R_m(A) \end{pmatrix}$$

where formula 6.36 has been applied. Therefore, multiplying A by $\mathcal{E}_{1,i}(\lambda)$ amounts

to multiplying the i -th row of A by λ . Similarly:

$$\mathcal{E}_{2;i,j}(\lambda) \cdot A = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_i + \lambda \epsilon_j \\ \vdots \\ \epsilon_m \end{pmatrix} \cdot A = \begin{pmatrix} \epsilon_1 \cdot A \\ \vdots \\ \epsilon_i \cdot A + \lambda \epsilon_j \cdot A \\ \vdots \\ \epsilon_m \cdot A \end{pmatrix} = \begin{pmatrix} R_1(A) \\ \vdots \\ R_i(A) + \lambda R_j(A) \\ \vdots \\ R_m(A) \end{pmatrix}.$$

This time, multiplying A by $\mathcal{E}_{2;i,j}(\lambda)$ amounts to adding to the i -th row of A the j -th row multiplied by λ . Finally,

$$\mathcal{E}_{3;i,j} \cdot A = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_j \\ \vdots \\ \epsilon_i \\ \vdots \\ \epsilon_m \end{pmatrix} \cdot A = \begin{pmatrix} \epsilon_1 \cdot A \\ \vdots \\ \epsilon_j \cdot A \\ \vdots \\ \epsilon_i \cdot A \\ \vdots \\ \epsilon_m \cdot A \end{pmatrix} = \begin{pmatrix} R_1(A) \\ \vdots \\ R_j(A) \\ \vdots \\ R_i(A) \\ \vdots \\ R_m(A) \end{pmatrix}$$

saying that multiplying A by $\mathcal{E}_{1,i}(\lambda)$ amounts in practice to exchanging the i -th row of A with the j -th row.

4.12.5 Esempio.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \end{pmatrix}$$

that is, the matrix $\mathcal{E}_{3;2,3}$ has exchanged the second row with the third.

4.12.6 Similarly, multiplying a matrix $A \in \mathbb{K}^{m \times n}$ on the right by elementary column matrices of type $\mathcal{F}$ amounts to performing column operations on the matrix A . Indeed:

$$\begin{aligned} A \cdot \mathcal{F}_{1;i}(\lambda) &= A \cdot (\mathbf{e}_1, \dots, \lambda \mathbf{e}_i, \dots, \mathbf{e}_n) = \\ &= (A \cdot \mathbf{e}_1, \dots, \lambda A \cdot \mathbf{e}_i, \dots, A \cdot \mathbf{e}_n) = (C_1(A), \dots, \lambda C_i(A), \dots, C_n(A)), \end{aligned}$$

that is, multiplication on the right by $\mathcal{F}_{1;i}(\lambda)$ has the effect of replacing the i -th column of A by its multiple by the scalar λ . The reader is left the undeniable task of verifying that $A \cdot \mathcal{F}_{2;i,j}(\lambda)$ is a matrix obtained from A by adding to its i -th column the j -th column multiplied by λ , whereas $A \cdot \mathcal{F}_{3;i,j}$ has the effect of exchanging columns i and j .

4.12.7 To sum up. An *elementary row transformation* on a matrix A is a transformation obtained by multiplying A on the left by a matrix $\mathcal{E}$. An *elementary column transformation* on a matrix A is a transformation obtained by multiplying A on the right by a matrix $\mathcal{F}$.

4.12.8 Esempio. Recall that multiplying a matrix A by $\mathbf{e}_j$ has the effect of returning the j -th column $C_j(A)$ of A . Let us now pretend that $\mathbf{e}_j$ is really an “operator” that returns the j -th column of a list of columns of arbitrary objects. Let then **Alph** be the set of letters of the alphabet and consider the word CANE, regarded as a row matrix with four columns. Then, “multiplying the word cane by the elementary matrix that exchanges the second and the fourth column yields:

$$\begin{aligned} (C \ A \ N \ E) \cdot (\mathbf{e}_1, \mathbf{e}_4, \mathbf{e}_3, \mathbf{e}_2) &= \\ &= ((C \ A \ N \ E)\mathbf{e}_1, (C \ A \ N \ E)\mathbf{e}_4, (C \ A \ N \ E)\mathbf{e}_3, (C \ A \ N \ E)\mathbf{e}_2) = \\ &= (C \ E \ N \ A) \end{aligned}$$

Conclusion: an anagram of CANE is CENA. And of Gatto?

The next proposition brings us the good news that every matrix $A \in \mathbb{K}^{m \times n}$ can be row-reduced by means of a finite number of elementary row transformations. More precisely:

4.12.9 Proposition. *For every matrix $A \in \mathbb{K}^{m \times n}$ there exists a finite number of elementary row matrices $E_1, E_2, \dots, E_p$ (that is: each $E_i \in \mathcal{E}$ denotes a suitable transformation of one of the three types listed) such that*

$$E_p \cdot \dots \cdot E_1 \cdot A$$

is a row-reduced matrix.

Proof. The proof will be carried out by induction on the number of rows. If A is a row, it is reduced. Indeed, either it is the zero row, and thus trivially satisfies the definition, or, if it is nonzero, every nonzero element is special. Therefore the proposition is true for all one-row matrices (that is, for all rows). Now suppose that the proposition is true for every matrix $A \in \mathbb{K}^{(m-1) \times n}$, for every $n \in \mathbb{N}$. Let us prove that it is true for every $A \in \mathbb{K}^{m \times n}$. Let then $A \in \mathbb{K}^{m \times n}$. If the matrix is zero, there is nothing to prove. If $A \neq \mathbf{0}$, it is no restriction to suppose that the first row is nonzero. If it were not, let i be the smallest integer such that $R_i(A) \neq 0$. Then the elementary row transformation $R_1(A) \leftrightarrow R_i(A)$, which exchanges the i -th row with the first, obtained by multiplying A by $\mathcal{E}_{3;1,i}$, yields a nonzero matrix whose first row

is nonzero. Let then $1 \leq j_0 \leq n$ be such that $A(1, j_0) \neq 0$. Up to multiplying the first row by

$$\lambda = \frac{1}{A(1, j_0)}$$

(by means of the transformation $\mathcal{E}_{1,1}(\lambda)$), one may assume that $(A)(1, j_0) = 1$. Let $R_2(A), \dots, R_m(A)$ be the other rows of the matrix A . One now performs, for every $2 \leq i \leq m$, the row transformations $R_i(A) \rightarrow R_i(A) - A(i, j_0)R_1(A)$ (that is, let us multiply the matrix A by $\mathcal{E}_{2;1,i}(A(i, j_0))$, for every $2 \leq i \leq m$). Such operations transform A into a matrix $A' \in \mathbb{K}^{m \times n}$ such that $A(1, j_0) = 1$ and $(A')(i, j_0) = 0$ for every $2 \leq i \leq m$, so that the element $A(i, j_0) = 1$ of A' is special. Hence A is row-reducible if and only if A' is row-reducible. But in order to continue reducing A' , it is enough to reduce the submatrix $A'' \in \mathbb{K}^{m-1 \times n-1}$ obtained from A' by “deleting” the first row and the j_0 -th column. Since by hypothesis a matrix with $m-1$ rows is row-reducible, then A itself is row-reducible, thus concluding the proof of the theorem. ■

The proof of the proposition explains how to row-reduce a matrix. Consider the following:

4.12.10 Esempio. To reduce the matrix

$$A := \begin{pmatrix} 2 & 3 & 6 & 9 & 8 \\ 1 & 3 & 5 & -1 & 3 \\ 4 & 3 & 0 & 0 & 3 \\ 2 & 2 & 5 & 3 & 1 \end{pmatrix}$$

one may begin by exchanging the first and the second row (solely for the purpose of simplifying the calculations, as will be seen shortly). Let then A' be the matrix such that $R_2(A') = R_2(A) - 2R_1(A)$, $R_3(A') = R_3(A) - 4R_1(A)$, $R_4(A') \rightarrow R_4(A) - 2R_1(A)$, obtaining:

$$A' := \begin{pmatrix} 1 & 3 & 5 & -1 & 3 \\ 0 & -3 & -4 & 11 & 2 \\ 0 & -9 & -20 & 4 & -9 \\ 0 & -4 & -5 & 5 & -5 \end{pmatrix}$$

To reduce the matrix A' it is therefore sufficient to reduce the matrix

$$B := \begin{pmatrix} -3 & -4 & 11 & 2 \\ -9 & -20 & 4 & -9 \\ -4 & -5 & 5 & -5 \end{pmatrix}$$

Perform the transformations $R_3(A'') = R_3(A') - 3R_2(A')$, $R_4(A') = 3R_4(A') - 4R_2(A')$.

One obtains

$$A'' := \begin{pmatrix} 1 & 3 & 5 & -1 & 3 \\ 0 & -3 & -4 & 11 & 2 \\ 0 & 0 & -8 & -29 & -15 \\ 0 & 0 & 1 & -29 & -23 \end{pmatrix}$$

Let now A''' be the matrix obtained by the transformation $R_4(A''') = R_4(A'') - R_3(A'')$. Then A''' is reduced and the special elements are highlighted by a little rectangle:

$$A''' = \begin{pmatrix} \boxed{1} & 3 & 5 & -1 & 3 \\ 0 & \boxed{-3} & -4 & 11 & 2 \\ 0 & 0 & -8 & \boxed{-29} & -15 \\ 0 & 0 & \boxed{9} & 0 & \boxed{-8} \end{pmatrix}$$

4.12.11 Esempio. To reduce the matrix A of example 6.12.10, one could have begun more simply by exploiting the zeros already available. For brevity, we shall limit ourselves to indicating in sequence the row operations involved and the final result. Let A' be obtained by the exchange $R_2(A) \leftrightarrow R_1(A)$, and let a matrix A'' be obtained by the exchange $R_2(A') \leftrightarrow R_3(A')$, obtaining:

$$A'' := \begin{pmatrix} 1 & 3 & 5 & -1 & 3 \\ 4 & 3 & 0 & 0 & 3 \\ 2 & 3 & 6 & 9 & 8 \\ 2 & 2 & 5 & 3 & 1 \end{pmatrix}$$

Let A''' be built using the transformations $R_3(A''') \rightarrow R_3(A'') + 9R_1(A'')$ and $R_4(A''') \rightarrow R_4(A'') + 3R_1(A'')$, obtaining:

$$A''' = \begin{pmatrix} 1 & 3 & 5 & -1 & 3 \\ 4 & 3 & 0 & 0 & 3 \\ 11 & 30 & 51 & 0 & 35 \\ 5 & 11 & 20 & 0 & 10 \end{pmatrix}$$

The completion of the calculations is entrusted to the reader. The new reduced form will not necessarily be equal to the one obtained in example 6.12.10.

An immediate consequence of Proposition 6.12.9 is therefore that every matrix can be reduced by columns. Indeed, its transpose can be reduced by rows.

4.12.12 Two matrices $A, B \in \mathbb{K}^{m \times n}$ are said to be row-equivalent if and only if there exist row elementary transformations $E_1, \dots, E_p$ such that

$$B := E_p \cdot \dots \cdot E_1 \cdot A.$$

They are said to be column-equivalent if and only if there exist $F_1, \dots, F_q$ such that

$$B = A \cdot F_1 \cdot \dots \cdot F_q.$$

4.12.13 Proposition. *The relation $A \sim B$ if and only if A and B are row-equivalent (resp. column-equivalent) is an equivalence relation.*

Proof. We shall prove the proposition for rows, leaving to the mirror of the reader the corresponding proof for columns. The relation is clearly reflexive, since $A = \mathbf{1}_m \cdot A$ and the identity is an elementary matrix. Moreover, since every elementary matrix is invertible, $A \sim B$ implies $B \sim A$. Finally, if $A \sim B$ and $B \sim C$, then there exist $E_1, \dots, E_p$ and $E'_1, \dots, E'_q$ such that

$$B = E_p \cdot \dots \cdot E_1 \cdot A \quad \text{and} \quad C = E'_q \cdot \dots \cdot E'_1 \cdot B.$$

hence

$$C = E'_q \cdot \dots \cdot E'_1 \cdot E_p \cdot \dots \cdot E_1 \cdot A,$$

whence $A \sim C$. ■

4.13 Remarks on the inverse of a square matrix.

4.13.1 Observe that every elementary matrix is invertible. Indeed it is easy to verify that

$$\mathcal{E}_{1,i}(\lambda)^{-1} = \mathcal{E}_{1,i}\left(\frac{1}{\lambda}\right), \quad \lambda \neq 0; \quad \mathcal{E}_{2;i,j}(\lambda)^{-1} = \mathcal{E}_{2;i,j}(-\lambda); \quad \mathcal{E}_{3;i,j}^{-1} = \mathcal{E}_{3;i,j}. \quad (4.61)$$

Similarly

$$\mathcal{F}_{1,i}(\lambda)^{-1} = \mathcal{F}_{1,i}\left(\frac{1}{\lambda}\right), \quad \lambda \neq 0; \quad \mathcal{F}_{2;i,j}(\lambda)^{-1} = \mathcal{F}_{2;i,j}(-\lambda); \quad \mathcal{F}_{3;i,j}^{-1} = \mathcal{F}_{3;i,j}.$$

There is no great science required to verify the preceding equalities. For example, the second of (6.61) says that if $A' = \mathcal{E}_{2;i,j}(\lambda) \cdot A$, then all the rows of A' coincide with those of A except the i -th one, which is:

$$R_i(A') = R_i(A) + \lambda \cdot R_j(A).$$

Therefore the matrix $\mathcal{E}_{2;i,j}(-\lambda)A'$ has all rows equal to those of the matrix A , except the i -th:

$$R_i(\mathcal{E}_{2;i,j}(-\lambda) \cdot A') = R_i(A') - \lambda R_j(A').$$

Now $R_j(A') = R_j(A)$, since $i \neq j$, and therefore:

$$R_i(\mathcal{E}_{2;i,j}(-\lambda) \cdot A') = R_i(A') - \lambda R_j(A) = R_i(A) + \lambda \cdot R_j(A) - \lambda R_j(A) = R_i(A).$$

Therefore $\mathcal{E}_{2;i,j}(-\lambda)A' = A$, that is, $\mathcal{E}_{2;i,j}(-\lambda)$ is the inverse of $\mathcal{E}_{2;i,j}(\lambda)$.

4.13.2 Question. The argument just presented correctly proves that the matrix $\mathcal{E}_{2;i,j}(\lambda)$ is invertible and that its inverse is $\mathcal{E}_{2;i,j}(-\lambda)$. One small detail, however, has been passed over. Is the reader able to identify it and take care of it?

4.13.3 Proposition. *Let A be an $n \times n$ square matrix. Then A is row-equivalent to a diagonal matrix, that is, there exists a finite sequence $E_1, \dots, E_h$ of elementary row transformations such that $E_h \dots E_2 E_1 \cdot A$ is a diagonal matrix.*

Proof. First of all there exist elementary row matrices $E_1, \dots, E_h$ such that $A_1 := E_h \dots E_2 E_1 \cdot A$ is row-reduced and upper triangular, that is, having only zeros below the main diagonal. Let $a_1, \dots, a_n$ be the elements of the main diagonal of A_1 . Let $C_j(A)$ be the first nonzero column. Up to an exchange of rows one may suppose that $A(j, j) \neq 0$ and therefore one may apply elementary row transformations so as to have $A(j, j)$ as a special element. One proceeds analogously with the first nonzero column after the j -th one. Continuing in this way, one obtains a matrix that is row-reduced and upper triangular. Now, with elementary row transformations of the form

$$R_i(A_1) \rightarrow R_i(A_1) - \frac{A(i, n)}{a_n} R_n(A_1), \quad 1 \leq i \leq n-1,$$

one obtains a matrix A_2 such that $C_n(A_2) = a_n \mathbf{e}_n$, that is, having only zeros above the element a_n . With the transformations

$$R_i(A_2) \rightarrow R_i(A_1) - \frac{A(i, n-1)}{a_{n-1}} R_n(A_1), \quad 1 \leq i \leq n-2$$

one has $C_{n-1}(A_2) = a_{n-1} \mathbf{e}_{n-1}$ and $C_n(A_2) = a_n \mathbf{e}_n$. By induction, continuing to apply elementary row transformations of the type mentioned, one obtains a diagonal matrix. ■

4.13.4 Esempio. Let us illustrate the proof of Proposition 6.13.3 on an example, in order to visualize how the algorithm works in practice. Let

$$A := \begin{pmatrix} 3 & 5 & 1 \\ 3 & 10 & 3 \\ 6 & 5 & 2 \end{pmatrix}.$$

The first step consists in row-reducing A . Let A_1 be the matrix such that

$$R_2(A_1) = R_2(A) - R_1(A) \quad \text{and} \quad R_3(A_1) = R_3(A) - 2R_1(A).$$

One obtains:

$$A_1 = \begin{pmatrix} 3 & 5 & 1 \\ 0 & 5 & 2 \\ 0 & -5 & 0 \end{pmatrix}.$$

Let now A_2 be the matrix having the same rows as A_1 except for the third, given by:

$$R_3(A_2) = R_3(A_1) + R_2(A_1).$$

which is already in triangular form.

$$A_2 = \begin{pmatrix} 3 & 5 & 1 \\ 0 & 5 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

Now A_2 is in triangular form. Let A_3 be such that

$$R_3(A_3) = R_3(A_2) \quad R_2(A_3) = R_2(A_2) - R_3(A_2) \quad R_1(A_3) = R_1(A_2) - \frac{1}{2}R_3(A_2)$$

Then:

$$A_3 = \begin{pmatrix} 3 & 5 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

With one further row transformation one obtains the matrix A_4 , having all rows equal to those of A_3 except the first, given by $R_1(A_4) \rightarrow R_1(A_3) - R_2(A_3)$, thus obtaining:

$$A_4 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

which is the desired diagonal matrix.

4.13.5 Exercise. Find a diagonal form row-equivalent to the matrix:

$$A := \begin{pmatrix} 0 & 1 & 0 & 3 \\ 3 & 1 & 0 & 1 \\ 6 & 2 & 0 & 4 \\ 9 & 1 & 0 & 3 \end{pmatrix}.$$

4.13.6 Corollary. *Every invertible $n \times n$ square matrix is row-equivalent to the $n \times n$ identity matrix.*

Proof. The matrix is invertible if and only if it has maximal rank. Obtain the diagonal form by means of elementary row transformations such as those described in Proposition 6.13.3. Then all the elements of the main diagonal, say $a_1, \dots, a_n$, are nonzero, since elementary transformations preserve rank, and if one of the diagonal elements were zero there would be a whole zero row (or column), and the rank would then be less than n . Multiplying A_1 by the matrices $\mathcal{E}_{1,i}(1/a_i)$, $1 \leq i \leq n$, one finally obtains the identity matrix, and this proves the proposition. ■

4.13.7 Corollary 6.13.6 suggests a practical method for computing the inverse of an invertible square matrix. Indeed let $E_1, \dots, E_h$ be respectively row and column elementary matrices such that

$$E_h \cdot \dots \cdot E_2 \cdot E_1 \cdot A = \mathbf{1}_n. \quad (4.62)$$

Then, by definition of inverse, $E_h \cdot \dots \cdot E_2 \cdot E_1$ is the inverse of A . Therefore:

$$A^{-1} = E_h \cdot \dots \cdot E_2 \cdot E_1.$$

Consider then the $n \times 2n$ matrix

$$(A \mid \mathbf{1}_n),$$

and let $E_1, \dots, E_h$ be elementary matrices satisfying (6.62). Then

$$\begin{aligned} E_h \cdot \dots \cdot E_2 \cdot E_1 \cdot (A \mid \mathbf{1}_n) &= (E_h \cdot \dots \cdot E_2 \cdot E_1 \cdot A \mid E_h \cdot \dots \cdot E_2 \cdot E_1 \mathbf{1}_n) = \\ &= (E_h \cdot \dots \cdot E_2 \cdot E_1 \cdot A \mid E_h \cdot \dots \cdot E_2 \cdot E_1) = (\mathbf{1}_n \mid E_h \cdot \dots \cdot E_2 \cdot E_1). \end{aligned}$$

This means that if one multiplies the matrix $(A \mid \mathbf{1}_n)$ by the same elementary matrices $E_1, \dots, E_h$ satisfying (6.62), the first n columns give the identity matrix, while the last n give the inverse.

4.13.8 Esempio. Suppose one wishes to find the inverse of the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

The matrix is invertible since $|A| = -3 \neq 0$. To find its inverse one considers the matrix

$$B = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array} \right)$$

Let B_1 be such that $R_1(B_1) = R_1(B)$ and $R_2(B_1) = R_2(B) - 2R_1(B)$, that is:

$$B_1 = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -3 & -2 & 1 \end{array} \right).$$

Let now B_2 be such that $R_2(B_2) = R_2(B_1)$ and $R_1(B_2) = R_1(B_1) + \frac{2}{3}R_2(B_1)$. Then:

$$B_2 = \left(\begin{array}{cc|cc} 1 & 0 & -\frac{1}{3} & \frac{2}{3} \\ 0 & -3 & -2 & 1 \end{array} \right).$$

Let finally B_3 be the matrix obtained from B_2 by multiplying the second row by $-1/3$. One has:

$$B_3 = \left(\begin{array}{cc|cc} 1 & 0 & -\frac{1}{3} & \frac{2}{3} \\ 0 & 1 & \frac{2}{3} & -\frac{1}{3} \end{array} \right).$$

Since the first two columns of B_3 form the 2×2 identity matrix, the last two are precisely the inverse of the matrix A , that is:

$$A^{-1} = -\frac{1}{3} \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$$

as we would have known how to determine by another route.

4.13.9 Since reducing a matrix by columns is essentially equivalent to reducing its transpose by rows, all the propositions mentioned can be reformulated in terms of columns. That is: every square matrix is column-equivalent to a diagonal matrix. Moreover: every invertible square matrix is column-equivalent to the identity matrix and, therefore, the inverse of a matrix may be found by an analogous procedure, multiplying the matrix $(A \mid B)$ on the right by matrices $F_1, \dots, F_h$ such that $A \cdot F_1 \cdot \dots \cdot F_h = \mathbf{1}_n$.