Chapter 4

Matrices

In the analysis of economic phenomena, we need to study the data that describe their behavior. Sometimes it may be possible to draw certain conclusions through simple visual examination, but in general, the figures and values taken to analyze a phenomenon or situation will hide the essential information that will be the basis of our analysis. Extracting this information is not always a simple task, and we then need the help of complex methods. Mathematics and statistics provide the key tools that allow revealing the internal structures of the data that are not accessible through direct observation. To bring this hidden information to light, it is necessary to apply sophisticated mathematical techniques, and in this chapter we introduce the basic mathematical structures for the representation and manipulation of information: tuples and matrices.

4.1 Basic Definitions

In the following definition, we present more precisely the concepts of tuple and matrix that we mentioned before.

Definition 86.

- Given $n \in \mathbb{N}$, we call an n-tuple any ordered set of n real numbers of the form

$$(a_1, a_2, \ldots, a_n).$$

  - The set of all n-tuples is denoted $\mathbb{R}^n$. Therefore,

$$\mathbb{R}^n = \{(a_1, a_2, \ldots, a_n) : a_1, a_2, \ldots, a_n \in \mathbb{R}\}.$$

  - The numbers $a_1, a_2, \ldots, a_n$ are called elements of the n-tuple.
  - Two n-tuples are equal if they have the same elements in the same positions. That is,

$$(a_1, a_2, \ldots, a_n) = (b_1, b_2, \ldots, b_n) \Leftrightarrow \begin{cases} a_1 = b_1, \\ a_2 = b_2, \\ \vdots, \\ a_n = b_n. \end{cases}$$

  - 2-tuples are called pairs. 3-tuples are called triples.

- We call a matrix of real numbers with m rows and n columns or of type $m \times n$ (or of order $m \times n$) a set of real numbers ordered in the form,

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

where for each $i = 1, \dots, m$ and $j = 1, \dots, n$, $a_{ij} \in \mathbb{R}$ is the number located in row i and column j. A matrix $m \times n$ is, therefore, a table or array of numbers with m rows and n columns.

- The numbers that make up the matrix are called elements or coefficients of the matrix.
- The (i, j) element of the matrix is the one found in row i and column j.
- Matrices are named using capital letters (A, B, C, etc.). Given a matrix, A, its elements are generically designated by the corresponding lowercase letter and the subscripts indicating the row and column. Thus the (i, j) element of matrix A will be a_{ij}.
- The set of all matrices of type $m \times n$ is denoted as $\mathcal{M}_{m\times n}$:

$$\mathcal{M}_{m\times n} = \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} / a_{ij} \in \mathbb{R}, \forall i, j \right\}.$$

  Given a matrix, A, we sometimes denote that it is of type $m \times n$ by writing $A_{m\times n}$.
- The generic matrix $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ is abbreviated by $(a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ or $(a_{ij})_{m\times n}$.
- Two matrices $A = (a_{ij})_{m\times n} \in \mathcal{M}_{m\times n}$ and $B = (b_{ij})_{\overline{m}\times\overline{n}} \in \mathcal{M}_{\overline{m}\times\overline{n}}$ are equal if it holds that:

$$\begin{cases} m = \overline{m} \\ n = \overline{n} \\ a_{ij} = b_{ij}, \forall i = 1, \dots, m, \forall j = 1, \dots, n \end{cases}.$$

  That is, if they are of the same type and have the same elements located in the same place.
- The n-tuple $v = (a_1, a_2, \dots, a_n)$ can be written in the form of a row matrix or column matrix as

$$v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \qquad \text{or} \qquad v = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}.$$

Examples 87.
1) On numerous occasions, a single number is sufficient to describe the state of a certain object or real-world situation. For example, if we are studying the profitability of a company in different years, in principle it will be sufficient to indicate the income it obtains at each moment and that income is expressed by a single number. Thus we will say:

- The first year the profits were 3 million euros.
- The second year the profits were 3.5 million euros.
- The third year the profits were 3.9 million euros.

In each case, the profitability is described by giving a single number (first 3, then 3.5 and finally 3.9).

However, a more detailed study of the company's profitability would require taking into account not only the final profits but also the volume of income and expenses. For example, a company with income of 4 million and expenses of 1 million will have profits of 3 million and high profitability. In contrast, a company that earns 103 million with expenses of 100 will have the same 3 million in profits but we should deduce that its profitability is much lower since to obtain equal amounts the expenses are much higher. In this way, if we take into account income and expenses, we will work with data of the type:

- The first year, the income was 13 million and the expenses were 10 million.

- The second year, the income was 14 million and the expenses were 10.5 million.

- The third year, the income was 14.3 million and the expenses were 10.4.

To abbreviate, we could agree to present the data for each year in an orderly manner in the form of a row or column as follows,

$$\boxed{\text{(income,expenses)}} \quad \text{or} \quad \boxed{\begin{pmatrix} \text{income} \\ \text{expenses} \end{pmatrix}}$$

In this way to indicate the profitability data in a certain period, it will be enough to write something like (15,11) and we will then know that we are referring to an income of 15 million euros and expenses of 11. For example, the same data as before can be written as:

- The first year the profitability data are (13,10) or $\begin{pmatrix} 13 \\ 10 \end{pmatrix}$.

- The second year the profitability data are (14,10.5) or $\begin{pmatrix} 14 \\ 10.5 \end{pmatrix}$.

- The third year the profitability data are (14.3,10.4) or $\begin{pmatrix} 14.3 \\ 10.4 \end{pmatrix}$.

It is evident that if we are going to write the information following this format, the order of the data is fundamental; thus, the pairs (15,11) and (11,15) represent different profitability data (note that the last pair corresponds to a company with higher expenses than income and therefore with losses). This justifies that we use an order when giving the data, that is, that we use an ordered pair.

We see then that we can represent through an ordered pair of real numbers the different possibilities of income and expenses that can occur. However, we know that a pair of real numbers is nothing more than a 2-tuple or element of $\mathbb{R}^2$. Therefore, in all of the above, what we do is to use elements of $\mathbb{R}^2$ to represent our information. Schematically the idea is the following:

Information of the problem		**Representation**
<u>Two</u> data: income and expenses of a company at a certain moment	$\Rightarrow$	<u>Two</u> real numbers ordered $\Downarrow$ element of $\mathbb{R}^{\underline{2}}$ 2-tuple

The need to pose mathematical models for phenomena in which several data intervene simultaneously is what motivates the use of elements of $\mathbb{R}^2$ (when we have two data), of $\mathbb{R}^3$ (when we have three) and in general of $\mathbb{R}^n$ (for n data).

It is even possible that we have data that require more complex structures. For example, following the previous example of income and expenses of companies, it could be that we had the information of three different companies for a specific year:

- Data of the first company: (13,10).

- Data of the second company: (19,15).

- Data of the third company: (17,12).

We can represent all these data together by arranging them orderly by rows in the form of a table,

$$
\begin{array}{l} \\ 1^{\text{st}} \text{ company} \\ 2^{\text{nd}} \text{ company} \\ 3^{\text{rd}} \text{ company} \end{array}
\quad
\begin{array}{cc} \text{income} & \text{expenses} \\ \end{array}
\begin{pmatrix} 13 & 10 \\ 19 & 15 \\ 17 & 12 \end{pmatrix},
$$

or we could also have written them in columns as

$$
\begin{array}{l} \\ \text{Income} \\ \text{Expenses} \end{array}
\quad
\begin{array}{ccc} 1^{\text{st}} \text{ company} & 2^{\text{nd}} \text{ company} & 3^{\text{rd}} \text{ company} \end{array}
\begin{pmatrix} 13 & 19 & 17 \\ 10 & 15 & 12 \end{pmatrix}.
$$

In both cases, the important thing is to fix a criterion for ordering the data. In both cases, the data ordered in the form of a table constitute what is called a matrix. Specifically,

- $\begin{pmatrix} 13 & 10 \\ 19 & 15 \\ 17 & 12 \end{pmatrix}$ is a matrix with three rows and two columns,
- $\begin{pmatrix} 13 & 19 & 17 \\ 10 & 15 & 12 \end{pmatrix}$ is a matrix with two rows and three columns.

The concept of a matrix is of great importance in mathematics since, together with that of a pair or n-tuple, they are two of the main methods for the ordered representation of data. We will see throughout the topic that matrices and elements of $\mathbb{R}^n$ allow modeling different real-life situations and natural phenomena.

2) Take the matrix

$$A = \left(\begin{array}{ccc} 3 & -1 & 2 \\ 0 & -12 & 4 \end{array} \right).$$

It is a matrix of order 2×3 since it has two rows and three columns. Furthermore,

- the $(2,1)$ element of A is 0,
- the $(2,3)$ element of A is 4,
- the $(3,1)$ element of A does not exist,
- etc.

Likewise, we have that A is an element of the set of all matrices of type 2×3, that is, $A \in \mathcal{M}_{2\times3}$.

3) Consider a matrix $B \in \mathcal{M}_{3\times2}$. Since B is an element of $\mathcal{M}_{3\times2}$, it will be a matrix with 3 rows and 2 columns. If we have no more information about B, we cannot know the values of its elements. In such a case, we must denote them in a generic way as indicated in Definition 86 using the lowercase corresponding to B (that is, b) and subscripts. Thus,

- the $(1,1)$ element of B will be $b_{1,1}$, - the $(1,2)$ element of B will be $b_{1,2}$,
- the $(2,1)$ element of B will be $b_{2,1}$, - the $(2,2)$ element of B will be $b_{2,2}$,
- the $(3,1)$ element of B will be $b_{3,1}$, - the $(3,2)$ element of B will be $b_{3,2}$.

In this way, placing each element in its place, the matrix B is

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \\ b_{3,1} & b_{3,2} \end{pmatrix}. \tag{4.1}$$

In example **2)**, since we know exactly what the matrix A is, we could indicate the value of each of its elements. However, in this example we do not know what real number appears in each position of B and therefore we are forced to assign it a generic name ($b_{2,1}$, $b_{3,1}$, etc.) that represents its value which we ignore.

Since writing the expression for B that appears in (4.1) repeatedly is tedious, instead we can abbreviate by writing $(b_{i,j})_{\substack{i=1,\dots,3 \\ j=1,2}}$ or even more briefly $(b_{i,j})_{3\times 2}$.

4) Let $A \in \mathcal{M}_{3\times 4}$. The matrix A will have 3 rows and 4 columns and if we do not know what the elements of A are, we must write them in a generic form as,

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \end{pmatrix}.$$

To abbreviate the writing we have that,

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \end{pmatrix} = \underbrace{(a_{i,j})_{\substack{i=1,\dots,3 \\ j=1,\dots,4}}}_{\text{the same more briefly}} = \underbrace{(a_{i,j})_{3\times 4}}_{\text{even more brief}}.$$

We can also indicate that A is a matrix with three rows and four columns by writing $A_{3\times 4}$.

5) The set $\mathcal{M}_{2\times 2}$ contains all matrices with two rows and two columns. All matrices in $\mathcal{M}_{2\times 2}$ are of the form

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

while the coefficients $a_{1,1}$, $a_{1,2}$, $a_{2,1}$, $a_{2,2}$ can take any value in $\mathbb{R}$. That is,

$$\mathcal{M}_{2\times 2} = \underbrace{\{ \underbrace{\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}}_{\substack{\text{Matrices with two} \\ \text{rows and two columns}}} : \underbrace{a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2} \in \mathbb{R}}_{\substack{\text{the coefficients take} \\ \text{any value in } \mathbb{R}}} \}}_{\text{All these matrices gathered form } \mathcal{M}_{2\times 2}}.$$

6) The matrices

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 0 & 6 & 2 \end{pmatrix} \text{ and } B = \begin{pmatrix} 2 & 3 & -1 \\ 0 & 6 & 2 \end{pmatrix}$$

have the same elements (the same real numbers 3, 2, -1, 0, 6 and 2 appear in them) but we see that they are not located in the same positions (in position $(1, 1)$, A has a 3 while in the same position B has a 2). Therefore both matrices are different, so $A \neq B$. If we now consider the matrix

$$C = \begin{pmatrix} 3 & 2 & -1 & 0 \\ 0 & 6 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

again A and C have the same elements and this time they even appear located in the same places. However, both matrices are of different types since A is of type 2×3 and C is 3×4. In that case we also cannot say that A and C are the same matrix and consequently $A \neq C$.

Remark. It is common to use mathematical notation to abbreviate writing. Thus, instead of writing

'take a matrix, A, with three rows and four columns'

we can put

'take $A \in \mathcal{M}_{3\times 4}$'

or

'take $A_{3\times 4}$'.

Veamos ahora, en la siguiente definición, una lista de conceptos básicos dentro de la teoría de matrices.

Definition 88 (Basic concepts about matrices)**.**

- Given $A = (a_{ij})_{m\times n}$, we call any matrix obtained by deleting rows and/or columns from A a submatrix of A.

Examples 89.

1) Given $A = \begin{pmatrix} 1 & 2 & 3 & 6 \\ -1 & 2 & 0 & 1 \\ 2 & 1 & 8 & 9 \end{pmatrix}$ we have that:

- $\begin{pmatrix} 2 & 3 & 6 \\ 2 & 0 & 1 \end{pmatrix}$ is a submatrix of A since it is obtained by deleting row 3 and column 1 of A:

column 1

$$\text{row 3} \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ -1 & 2 & 0 & 1 \\ 2 & 1 & 8 & 9 \end{pmatrix}.$$

- $\begin{pmatrix} 1 & 6 \\ 2 & 9 \end{pmatrix}$ is a submatrix of A since it is obtained by deleting row 2 and columns 2 and 3 of A:

column 2, column 3

$$\text{row 2} \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ -1 & 2 & 0 & 1 \\ 2 & 1 & 8 & 9 \end{pmatrix}.$$

- $\begin{pmatrix} 3 \end{pmatrix}$ is a submatrix of A since it is obtained by deleting rows 2 and 3 and columns 1, 2 and 4 of A:

column 1, column 2, column 4

$$\begin{matrix} \\ \text{row 2} \\ \text{row 3} \end{matrix} \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ -1 & 2 & 0 & 1 \\ 2 & 1 & 8 & 9 \end{pmatrix}.$$

- $\begin{pmatrix} 1 & 2 & 0 \\ -1 & 2 & 8 \end{pmatrix}$ is not a submatrix of A since it cannot be obtained by deleting complete rows and columns from A as can be seen by observing the arrangement of its elements within A:

$$\begin{pmatrix} \boxed{1} & \boxed{2} & 3 & 6 \\ \boxed{-1} & \boxed{2} & \boxed{0} & 1 \\ 2 & 1 & \boxed{8} & 9 \end{pmatrix}$$

2) Consider the matrix we obtained on page 114 when compiling the income and expense data of several companies:

$$\begin{array}{r} \\ 1^{\text{st}} \text{ company} \\ 2^{\text{nd}} \text{ company} \\ 3^{\text{rd}} \text{ company} \end{array} \begin{array}{c} \text{income} \quad \text{expenses} \\ \begin{pmatrix} 13 & 10 \\ 19 & 15 \\ 17 & 12 \end{pmatrix} \end{array}.$$

Suppose that for some reason we want to remove the second company from the study. The matrix corresponding to the new situation will be,

$$\begin{array}{r} \\ 1^{\text{st}} \text{ company} \\ 3^{\text{rd}} \text{ company} \end{array} \begin{array}{c} \text{income} \quad \text{expenses} \\ \begin{pmatrix} 13 & 10 \\ 17 & 12 \end{pmatrix} \end{array}$$

It is evident that the new matrix is a submatrix of the initial matrix since it is obtained by deleting its second row. Suppose that we further decide to restrict our study even more and only take into account the expenses of the companies. In that case the data matrix will be,

$$\begin{array}{r} \\ 1^{\text{st}} \text{ company} \\ 3^{\text{rd}} \text{ company} \end{array} \begin{array}{c} \text{expenses} \\ \begin{pmatrix} 10 \\ 12 \end{pmatrix} \end{array}$$

Again we have obtained a submatrix of the initial matrix since we have deleted complete rows and columns:

$$\begin{pmatrix} 13 & 10 \\ 19 & 15 \\ 17 & 12 \end{pmatrix} \xrightarrow[\text{delete row 2}]{} \begin{pmatrix} 13 & 10 \\ 19 & 15 \\ 17 & 12 \end{pmatrix} \longrightarrow \begin{pmatrix} 13 & 10 \\ 17 & 12 \end{pmatrix}$$

$$\xrightarrow[\text{delete row 2 and column 1}]{} \begin{pmatrix} 13 & 10 \\ 19 & 15 \\ 17 & 12 \end{pmatrix} \longrightarrow \begin{pmatrix} 10 \\ 12 \end{pmatrix}$$

- A matrix of the form $\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}_{1\times n}$ of type $1 \times n$ that has only one row is called a row matrix.

Example 90. The matrices $\begin{pmatrix} 2 & -1 & 0 & 1 \end{pmatrix}_{1\times 4}$, $\begin{pmatrix} -1 & 4 & 12 \end{pmatrix}_{1\times 3}$ or $\begin{pmatrix} 2 & 4 \end{pmatrix}_{1\times 2}$ are row matrices.

- A matrix of the form $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}_{n\times 1}$ of type $n \times 1$ that has only one column is called a column matrix.

Example 91. The matrices $\begin{pmatrix} 1 \\ 0 \\ -3 \\ 4 \end{pmatrix}_{4\times 1}$ and $\begin{pmatrix} 2 \\ 6 \end{pmatrix}_{2\times 1}$ are column matrices.

- Given the n-tuples $v_1, v_2, \ldots, v_n \in \mathbb{R}^m$:
  - The matrix obtained by grouping v_1, v_2, $\ldots$, v_n by columns is denoted

$$\begin{pmatrix} v_1 & | & v_2 & | & \cdots & | & v_n \end{pmatrix}$$

    and will have m rows and n columns. It will therefore be a matrix of $\mathcal{M}_{m\times n}$.
  - The matrix obtained by grouping v_1, v_2, $\ldots$, v_n by rows is denoted

$$\begin{pmatrix} v_1 \\ \hline v_2 \\ \hline \vdots \\ \hline v_n \end{pmatrix}$$

    which will have n rows and m columns. That is, it is of type $n \times m$.

Example 92. Let $v_1 = (2, 3, -1, 0)$, $v_2 = (6, 2, 3, 3)$, $v_3 = (6, 4, -9, -1)$. The block matrix obtained by grouping v_1, v_2 and v_3 by columns is

$$\begin{pmatrix} v_1 & | & v_2 & | & v_3 \end{pmatrix} = \begin{pmatrix} 2 & 6 & 6 \\ 3 & 2 & 4 \\ -1 & 3 & -9 \\ 0 & 3 & -1 \end{pmatrix} \in \mathcal{M}_{4\times 3}.$$

The block matrix we obtain by grouping v_1, v_2 and v_3 by rows is

$$\begin{pmatrix} v_1 \\ \hline v_2 \\ \hline v_3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & -1 & 0 \\ 6 & 2 & 3 & 3 \\ 6 & 4 & -9 & -1 \end{pmatrix} \in \mathcal{M}_{3\times 4}.$$

- A matrix with n rows and n columns (of type $n \times n$) is said to be a square matrix of order n. The set of all square matrices of order n is denoted by $\mathcal{M}_n$.

Example 93. The matrix $A = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix}$ has two rows and two columns. Therefore it is a square matrix of order 2. We know that the set of matrices of type 2×2 with two rows and two columns is denoted $M_{2\times 2}$. However, to abbreviate, we write $\mathcal{M}_2$ instead of $\mathcal{M}_{2\times 2}$. Therefore,

$$A \in \mathcal{M}_2 = \mathcal{M}_{2\times 2}.$$

In the same way, the matrix

$$B = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 8 & 2 \\ 7 & 6 & 5 \end{pmatrix}$$

has three rows and three columns and is therefore a square matrix of order 3 so $B \in \mathcal{M}_3 = \mathcal{M}_{3\times 3}$.

- Given $A \in \mathcal{M}_{m\times n}$ we call the transpose of A and denote it by A^t, the matrix whose first row is the first column of A, whose second row is the second column of A, ..., whose n-th row is the n-th column of A. The following facts are evident:

  - $A \in \mathcal{M}_{m\times n} \Rightarrow A^t \in \mathcal{M}_{n\times m}$.
  - $(A^t)^t = A$.
  - $A = (a_{ij})_{m\times n} \Rightarrow A^t = (a_{ji})_{n\times m}$. That is, the element located in A at position (i,j) when transposed moves to position (j,i).

Examples 94.

1) It is easy to calculate the transpose of any matrix. For example, the transpose of $\begin{pmatrix} 2 & 3 & 7 \\ 1 & 6 & 4 \end{pmatrix}$ is denoted $\begin{pmatrix} 2 & 3 & 7 \\ 1 & 6 & 4 \end{pmatrix}^t$ and is calculated by exchanging rows for columns according to the scheme,

column 1 column 2 column 3

$$\begin{pmatrix} 2 & 3 & 7 \\ 1 & 6 & 4 \end{pmatrix}^t = \begin{pmatrix} 2 & 1 \\ 3 & 6 \\ 7 & 4 \end{pmatrix} \begin{matrix} \text{row 1} \\ \text{row 2} \\ \text{row 3} \end{matrix} .$$

See that the initial matrix is of type 2×3 and when transposing we obtain one of type 3×2. It is also evident that if we transpose the last matrix again we will get the first one back:

$$\begin{pmatrix} 2 & 3 & 7 \\ 1 & 6 & 4 \end{pmatrix}^{tt} = \begin{pmatrix} 2 & 1 \\ 3 & 6 \\ 7 & 4 \end{pmatrix}^t = \begin{pmatrix} 2 & 3 & 7 \\ 1 & 6 & 4 \end{pmatrix}.$$

That is, doing the transpose twice is equivalent to doing nothing.

2) On page 114 we saw how we could organize in a matrix the data corresponding to the profitability of different companies. When writing the data for each of them we could arrange them by rows or by columns. It is evident that the matrix obtained by arranging the data by columns is the transpose of the one that appears when doing it by rows:

$$\begin{pmatrix} 13 & 10 \\ 19 & 15 \\ 17 & 12 \end{pmatrix}^t = \begin{pmatrix} 13 & 19 & 17 \\ 10 & 15 & 12 \end{pmatrix}.$$

Actually, a matrix and its transpose contain the same information but organized by columns instead of by rows or vice versa.

- We call the zero matrix of type $m \times n$ and denote it by $0_{m\times n}$ or simply 0, the matrix:

$$0_{m\times n} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}_{m\times n} \in \mathcal{M}_{m\times n}.$$

That is, the matrix $0_{m\times n}$ is the matrix of type $m \times n$ that has all its elements equal to zero.

Example 95. The zero matrix of type 2×4 is denoted $0_{2\times 4}$ and its value is

$$0_{2\times 4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

It is evident that $0_{2\times 4} \in \mathcal{M}_{2\times 4}$.

The zero matrix of type 3×3 is denoted $0_{3\times 3}$ and its value is

$$0_{3\times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

It is a matrix with three rows and three columns and therefore it is a square matrix that belongs to $\mathcal{M}_{3\times 3} = \mathcal{M}_3$.

In general we can consider the zero matrix of any type. For example,

$$0_{3\times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad 0_{2\times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad 0_{4\times 1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

It should be noted that sometimes, to abbreviate writing, the zero matrix is usually denoted simply by writing '0' (without indicating the type), however, this can lead to confusion since we see that there is not a single zero matrix but infinitely many of them, one for each type. In those cases where the zero matrix is denoted by '0', we must deduce what the type is from the context in which it appears.

Square matrices play a particularly important role in matrix theory and in matrix mathematical models. In the following definition we will see a list of concepts and definitions all of them relative to square matrices.

Definition 96 (Basic concepts for square matrices)**.**

- We call the main diagonal of the matrix $A_{n\times n} = (a_{ij})_{n\times n} \in \mathcal{M}_n$ the row matrix $\begin{pmatrix} a_{11} & a_{22} & \dots & a_{nn} \end{pmatrix}$. The main diagonal is therefore the row matrix formed by the elements of A that are boxed in the following representation:

$$A = \begin{pmatrix} \boxed{a_{11}} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & \boxed{a_{22}} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & \boxed{a_{33}} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & \boxed{a_{nn}} \end{pmatrix}.$$

- We call the trace of $A = (a_{ij})_{n\times n} \in \mathcal{M}_n$ and denote it by $tr(A)$ or $trace(A)$ the sum of the elements of the main diagonal:

$$tr(A) = a_{11} + a_{22} + \dots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

Example 97. Let $A = \begin{pmatrix} 1 & 2 & -6 \\ -5 & 2 & 4 \\ 3 & 7 & 9 \end{pmatrix}$ then we have that

- the main diagonal of A is $\begin{pmatrix} 1 & 2 & 9 \end{pmatrix}$.
- the trace of A is $\text{tr}(A) = 1 + 2 + 9 = 12$.

- We say that $(a_{ij})_{n\times n} \in \mathcal{M}_n$ is:
  - **upper triangular** if all the elements below the main diagonal are zero,
  - **lower triangular** if all the elements above the main diagonal are zero,
  - **diagonal** if all the elements outside the main diagonal are zero.

Examples 98.

1) $\begin{pmatrix} 1 & 2 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ are upper triangular matrices.

2) $\begin{pmatrix} 3 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 6 & 2 & -1 & 0 \\ 1 & 3 & 2 & 4 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ are lower triangular matrices.

3) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ are diagonal matrices.

- We call the identity matrix of order n and denote it by I_n the square matrix of order n that is diagonal and such that all the elements of its main diagonal are equal to 1. That is,

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \in \mathcal{M}_n.$$

Example 99.

- The identity matrix of order 1 is $I_1 = (1)$,
- the identity matrix of order 2 is $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$,
- the identity matrix of order 3 is $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

- We say that $(a_{ij})_{n\times n}$ is symmetric if $A = A^t$ or, equivalently:

$$a_{ij} = a_{ji}, \qquad \forall i, j = 1, \dots, n.$$

- We say that $(a_{ij})_{n\times n}$ is antisymmetric if it satisfies that:

$$a_{ij} = -a_{ji}, \qquad \forall i, j = 1, \dots, n.$$

Note that if the above holds, taking $i = j$ we obtain that

$$a_{ii} = -a_{ii} \Rightarrow 2a_{ii} = 0 \Rightarrow a_{ii} = 0, \quad \forall i = 1, \dots, n$$

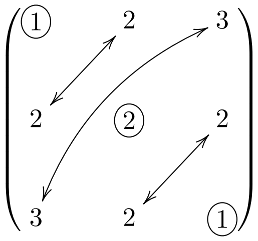
so an antisymmetric matrix will have all the elements of its main diagonal zero.

Examples 100.

1) Given $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 2 \\ 3 & 2 & 1 \end{pmatrix}$ we have that

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

and therefore A is a symmetric matrix. See in the following scheme that the elements (i, j) and (j, i) of the matrix coincide:

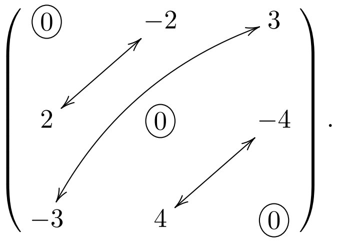


For the matrix to be symmetric, the elements located at the ends of the same arrow must coincide. Note that the elements of the main diagonal (circled) are not marked by any arrow and therefore their value does not influence in any way whether the matrix is symmetric.

2) The matrix $B = \begin{pmatrix} 0 & -2 & 3 \\ 2 & 0 & -4 \\ -3 & 4 & 0 \end{pmatrix}$ is antisymmetric since, if we denote $B = (b_{ij})_{3\times 3}$, we have that

* $b_{12} = -2 = -b_{21}$.
* $b_{13} = 3 = -b_{31}$.
* $b_{23} = -4 = -b_{32}$.
* $b_{11} = 0, b_{22} = 0, b_{33} = 0$.

The above can be schematized by the following arrow diagram:



The elements located at the ends of the same arrow must be opposites and those on the diagonal (enclosed in a circle) must be zero.

3) The matrix $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$ is not antisymmetric since the diagonal elements are not all zero. However the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is antisymmetric.

4) A supplier provides three different products A, B and C in different locations. It is observed that:

- In 3 locations only A is supplied.

- In 5 locations only B is supplied.
- In 4 locations only C is supplied.
- In 19 locations both A and B are supplied.
- In 23 locations both B and C are supplied.
- In 16 locations both A and C are supplied.

We can encode this data in a matrix in the following way,

$$\begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 3 & 19 & 16 \\ 19 & 5 & 23 \\ 16 & 23 & 4 \end{pmatrix} \end{array}.$$

We have placed at the intersection of row A with column A, the number of locations where only A is supplied, at the intersection of row A with column B, the locations where both A and B are supplied and so on. Evidently there are the same number of locations where A and B are sold as where B and A are distributed, so in the corresponding positions we put the same number, 19, and the same will happen in the case B-C and A-C. Therefore, the resulting matrix is symmetric since in position (i, j) we will always find the same number as in (j, i).

5) Three financial entities make loans among themselves so that one owes a certain amount of money to another:

- Entity 1 owes entity 2, 0.5 million euros,
- Entity 1 owes entity 3, 0.9 million euros,
- Entity 2 owes entity 3, −0.2 million euros.

It is evident that if entity 1 owes entity 2, 0.5 million euros, we can admit that entity 2 owes entity 1, −0.5 million. Thus, when we indicate that entity 2 owes entity 3, −0.2 million euros, we mean that the account status between both is favorable to entity 2 so that entity 3 owes it 0.2 million. On the other hand, it is also clear that each entity does not owe anything to itself. All this information we can summarize in the following matrix:

$$\text{The} \left\{ \begin{array}{l} \text{entity 1} \\ \text{entity 2} \\ \text{entity 3} \end{array} \right. \quad \begin{array}{c} \overbrace{\begin{array}{ccc} \text{entity 1} & \text{entity 2} & \text{entity 3} \end{array}}^{\text{owes to}} \\ \begin{pmatrix} 0 & 0.5 & 0.9 \\ -0.5 & 0 & -0.2 \\ -0.9 & 0.2 & 0 \end{pmatrix} \end{array}.$$

It can be seen that we have placed at the intersection of the rows and columns corresponding to the different entities the amount owed taking into account when putting the sign of each data that the rows indicate the debtor entity and the columns the creditor entity. As we said before, if one entity owes a certain amount to another, we can also say that the latter owes the former the same amount but with the opposite sign; therefore in the matrix we always find in position (i, j) the number opposite to the one that appears in (j, i). Also, since no entity owes anything to itself, on the main diagonal we have written only zeros. In short, the data matrix is antisymmetric.

Remark. Diagonal matrices are usually denoted, in order to abbreviate writing, by indicating only the elements of their main diagonal. Thus for example:

$$\text{The matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \text{ can be written as } \begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix}.$$

In generic form, the diagonal matrix $A \in \mathcal{M}_n$ whose main diagonal is $\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$ will be denoted by

$$A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}_{n\times n}.$$

4.2 Operations with matrices

We know that it is possible to perform different operations between real numbers. Thus, we can calculate the sum, difference, product, division, etc. We will see in this section that it is possible to extend these operations to matrix calculus. We will even verify that many of the usual properties of number arithmetic remain valid for matrices.

4.2.1 Matrix addition

Definition 101. Given two matrices of the same type $A = (a_{ij})_{m\times n}, B = (b_{ij})_{m\times n} \in \mathcal{M}_{m\times n}$ we define the sum of A and B as the matrix $A + B \in \mathcal{M}_{m\times n}$ determined by:

$$A + B = (a_{ij} + b_{ij})_{m\times n}.$$

That is:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}$$
$$= \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix}.$$

Examples 102.

1) $\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 2 & 5 \end{pmatrix}.$

2) $\begin{pmatrix} 1 & 2 & 0 & 9 \\ -1 & 4 & 6 & 6 \\ 2 & 1 & 2 & 4 \end{pmatrix} + \underbrace{\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}_{=0_{3\times 4}} = \begin{pmatrix} 1 & 2 & 0 & 9 \\ -1 & 4 & 6 & 6 \\ 2 & 1 & 2 & 4 \end{pmatrix}.$

3) $\begin{pmatrix} 1 & 2 & 0 \\ -1 & 4 & 6 \\ 2 & 1 & 2 \end{pmatrix} + \begin{pmatrix} -1 & -2 & 0 \\ 1 & -4 & -6 \\ -2 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0_{3\times 3}.$

4) $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 4 & 6 \end{pmatrix}$ is an operation that cannot be performed because the two matrices are not of the same type.

5) Since we have identified the elements of $\mathbb{R}^n$ with row and column matrices, the sum operation we have defined is also valid for them. Thus we can perform the following calculations:

- $(2,3,-1)+(6,0,2)=(8,3,1)$.
- $(3,-2)-(-3,2)=(0,0)$.
- $(4,3,2,1,1)+(-2,4,3,0,-1)=(2,7,5,1,0)$.
- $(3,2,1)+(2,4,6,2)$ is an operation that cannot be performed since we have different types.

Note that only matrices of the same type can be added and the result will then in turn be a matrix of that same type.

We compile below several properties of matrix addition. All of them are analogous to those that hold for the sum of two numbers and in fact their proof is derived directly from them.

Properties 103. *$\forall A,B,C\in\mathcal{M}_{m\times n}$:*

1. *Commutative property: $A+B=B+A$.*
2. *Associative property: $A+(B+C)=(A+B)+C$.*
3. *$A+0=A$ (where $0=0_{m\times n}$).*
4. *Given $A=(a_{ij})_{m\times n}$ we define the opposite matrix of A as*
$$-A=(-a_{ij})_{m\times n}\in\mathcal{M}_{m\times n}$$
and then it is verified that:
$$A+(-A)=0.$$
5. *$(A+B)^t=A^t+B^t$.*
6. *A is antisymmetric $\Leftrightarrow A^t=-A$*

Remark (Matrix subtraction)**.** In the same way that the sum is defined, it is equally easy to introduce matrix subtraction. In fact, as we see below, we can define subtraction from the sum.

Given two matrices $A=(a_{ij})_{m\times n}, B=(b_{ij})_{m\times n}\in\mathcal{M}_{m\times n}$ we will define the subtraction or difference between A and B as the matrix
$$A-B=A+(-B)=(a_{ij}-b_{ij})_{m\times n}\in\mathcal{M}_{m\times n}.$$

Examples 104.

1) $-\begin{pmatrix}1&2&1\\3&1&3\end{pmatrix}=\begin{pmatrix}-1&-2&-1\\-3&-1&-3\end{pmatrix}$.

2) $-0_{2\times 2}=-\begin{pmatrix}0&0\\0&0\end{pmatrix}=\begin{pmatrix}-0&-0\\-0&-0\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}=0_{2\times 2}$.

3) $\begin{pmatrix}1&2&4\\2&0&6\end{pmatrix}-\begin{pmatrix}2&-1&0\\3&2&3\end{pmatrix}=\begin{pmatrix}-1&3&4\\-1&-2&3\end{pmatrix}$.

4) $-(3,2,-1)=(-3,-2,1)$.

5) If $A = \begin{pmatrix} 0 & 7 \\ -7 & 0 \end{pmatrix}$ then, $A^t = \begin{pmatrix} 0 & -7 \\ 7 & 0 \end{pmatrix} = -\begin{pmatrix} 0 & 7 \\ -7 & 0 \end{pmatrix} = -A$, therefore A is an antisymmetric matrix.

4.2.2 Product of matrices by a real number

Definition 105. Given a matrix $A = (a_{ij})_{m\times n} \in \mathcal{M}_{m\times n}$ and a real number $r \in \mathbb{R}$, we define the product of r by A and denote it as $r \cdot A$ or $A \cdot r$ as:

$$r \cdot A = A \cdot r = (r \cdot a_{ij})_{m\times n} \in \mathcal{M}_{m\times n}.$$

That is:

$$r \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} r \cdot a_{11} & r \cdot a_{12} & \dots & r \cdot a_{1n} \\ r \cdot a_{21} & r \cdot a_{22} & \dots & r \cdot a_{2n} \\ \vdots & \vdots & & \vdots \\ r \cdot a_{m1} & r \cdot a_{m2} & \dots & r \cdot a_{mn} \end{pmatrix}.$$

Note that if we multiply a matrix of type $m \times n$ by a real number we obtain as a result a matrix of type $m \times n$.

Examples 106.
1) $2 \cdot \begin{pmatrix} 1 & 3 & 1 & 1 \\ 2 & 6 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 6 & 2 & 2 \\ 4 & 12 & 0 & -2 \end{pmatrix}.$

2) $8 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = 0_{3\times 2}.$

3) $-3 \cdot \begin{pmatrix} 1 & 2 \\ -1 & 6 \end{pmatrix} = \begin{pmatrix} -3 & -6 \\ 3 & -18 \end{pmatrix}.$

Again we summarize the most important properties of the product of a number by a matrix. In all cases the proof is evident and can be verified directly.

Properties 107. $\forall r, s \in \mathbb{R},\ \forall A, B \in \mathcal{M}_{m\times n}$:

1. *Distributive property:* $r \cdot (A + B) = r \cdot A + r \cdot B$.
2. *Distributive property:* $(r + s) \cdot A = r \cdot A + s \cdot A$.
3. $1 \cdot A = A$.
4. $(-1) \cdot A = -A$.
5. $(r \cdot s) \cdot A = r \cdot (s \cdot A)$.
6. $r \cdot 0 = 0$, $0 \cdot A = 0$.
7. $(r \cdot A)^t = r \cdot A^t$.

4.2.3 Product of two matrices

The operations of sum and product by a real number have a simple definition since it is sufficient to perform the corresponding operation, element by element. The product of matrices does not have such a direct formulation but we will see that the definition we give below is useful in different contexts and matrix models.

Definition 108. Given $A = (a_{ij})_{m\times n} \in \mathcal{M}_{m\times n}$ and $B = (b_{ij})_{n\times p} \in \mathcal{M}_{n\times p}$ we define the product of A and B and denote it $A \cdot B$, as the matrix

$$A \cdot B = (a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj})_{\substack{i=1,\ldots,m\\ j=1,\ldots,p}} \in \mathcal{M}_{m\times p}.$$

That is, in the position (i, j) of the matrix $A \cdot B$ is the element

$$\overbrace{a_{i1}\quad b_{1j} \quad + \quad a_{i2} \quad b_{2j} \quad + \quad a_{i3} \quad b_{3j} \quad + \quad \ldots \quad + \quad a_{in} \quad b_{nj}}^{\text{row } i \text{ of } A}\ ,$$
$$\underbrace{\phantom{a_{i1}\, b_{1j} + a_{i2}\, b_{2j} + a_{i3}\, b_{3j} + \ldots + a_{in}\, b_{nj}}}_{\text{column } j \text{ of } B}$$

which is obtained as the product of row i of A by column j of B.

Examples 109.

1) $\begin{pmatrix} 2 & 1 & 2 \\ 0 & -1 & 0 \end{pmatrix}_{2\times 3} \cdot \begin{pmatrix} 2 & 1 \\ -1 & 1 \\ 3 & 1 \end{pmatrix}_{3\times 2} =$

$$= \begin{pmatrix} \underbrace{\underbrace{\begin{pmatrix} 2 & 1 & 2 \end{pmatrix}}_{\text{row 1}} \cdot \underbrace{\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}}_{\text{column 1}}}_{=2\cdot2+1\cdot(-1)+2\cdot3=9} & \underbrace{\underbrace{\begin{pmatrix} 2 & 1 & 2 \end{pmatrix}}_{\text{row 1}} \cdot \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{\text{column 2}}}_{=2\cdot1+1\cdot1+2\cdot1=5} \\ \underbrace{\underbrace{\begin{pmatrix} 0 & -1 & 0 \end{pmatrix}}_{\text{row 2}} \cdot \underbrace{\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}}_{\text{column 1}}}_{=0\cdot2+(-1)\cdot(-1)+0\cdot3=1} & \underbrace{\underbrace{\begin{pmatrix} 0 & -1 & 0 \end{pmatrix}}_{\text{row 2}} \cdot \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{\text{column 2}}}_{=0\cdot1+(-1)\cdot1+0\cdot1=-1} \end{pmatrix}_{2\times 2} = \begin{pmatrix} 9 & 5 \\ 1 & -1 \end{pmatrix}_{2\times 2}.$$

2) $\begin{pmatrix} 1 & 3 & 6 \\ 3 & 9 & -1 \end{pmatrix}_{2\times 3} \cdot I_3 = \begin{pmatrix} 1 & 3 & 6 \\ 3 & 9 & -1 \end{pmatrix}_{2\times 3} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}_{3\times 3} =$

$$= \begin{pmatrix} \underbrace{\underbrace{\begin{pmatrix} 1 & 3 & 6 \end{pmatrix}}_{\text{row 1}} \cdot \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\text{column 1}}}_{=1\cdot1+3\cdot0+6\cdot0=1} & \underbrace{\underbrace{\begin{pmatrix} 1 & 3 & 6 \end{pmatrix}}_{\text{row 1}} \cdot \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\text{column 2}}}_{=1\cdot0+3\cdot1+6\cdot0=3} & \underbrace{\underbrace{\begin{pmatrix} 1 & 3 & 6 \end{pmatrix}}_{\text{row 1}} \cdot \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{\text{column 3}}}_{=1\cdot0+3\cdot0+6\cdot1=6} \\ \underbrace{\underbrace{\begin{pmatrix} 3 & 9 & -1 \end{pmatrix}}_{\text{row 2}} \cdot \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\text{column 1}}}_{=3\cdot1+9\cdot0+(-1)\cdot0=3} & \underbrace{\underbrace{\begin{pmatrix} 3 & 9 & -1 \end{pmatrix}}_{\text{row 2}} \cdot \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\text{column 2}}}_{=3\cdot0+9\cdot1+(-1)\cdot0=9} & \underbrace{\underbrace{\begin{pmatrix} 3 & 9 & -1 \end{pmatrix}}_{\text{row 2}} \cdot \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{\text{column 3}}}_{=3\cdot0+9\cdot0+(-1)\cdot1=-1} \end{pmatrix}_{2\times 3} = \begin{pmatrix} 1 & 3 & 6 \\ 3 & 9 & -1 \end{pmatrix}_{2\times 3}.$$

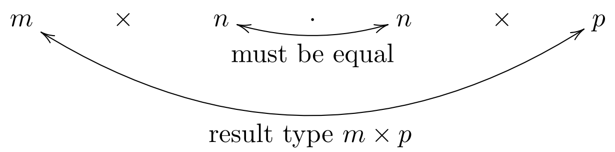
3) $\underbrace{\begin{pmatrix} 0 & 0 & 0 \end{pmatrix}}_{0_{1\times 3}} \cdot \begin{pmatrix} 2 & 2 \\ 1 & -1 \\ 3 & 2 \end{pmatrix}_{3\times 2} = \begin{pmatrix} 0 & 0 \end{pmatrix}_{1\times 2}.$

4) $\begin{pmatrix} 1 \\ 9 \\ 1 \end{pmatrix}_{3\times 1} \cdot \begin{pmatrix} 1 & 3 & 3 \end{pmatrix}_{1\times 3} = \begin{pmatrix} 1 & 3 & 3 \\ 9 & 27 & 27 \\ 1 & 3 & 3 \end{pmatrix}_{3\times 3}.$

5) $\begin{pmatrix} 1 & 3 & 3 \end{pmatrix}_{1\times 3} \cdot \begin{pmatrix} 1 \\ 9 \\ 1 \end{pmatrix}_{3\times 1} = \begin{pmatrix} 1\cdot 1 + 3\cdot 9 + 3\cdot 1 \end{pmatrix}_{1\times 1} = \begin{pmatrix} 31 \end{pmatrix}_{1\times 1} \equiv 31.$

6) $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}_{2\times 2} \cdot \begin{pmatrix} 1 & 3 \\ 2 & 1 \\ 3 & -1 \end{pmatrix}_{3\times 2}$ is an operation that cannot be performed because the types of matrices do not match.

Remark. Note that the product of a matrix of type $m \times n$ by a matrix $n \times p$ provides a matrix $m \times p$. Schematically we have:



Properties 110.

1. $\forall A \in \mathcal{M}_{m\times n}, B \in \mathcal{M}_{n\times p}, C \in \mathcal{M}_{p\times r}$

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

2. $\forall A \in \mathcal{M}_{m\times n}, B \in \mathcal{M}_{n\times p}, r \in \mathbb{R}$

$$A \cdot (r \cdot B) = (r \cdot A) \cdot B = r \cdot (A \cdot B).$$

3. $\forall A, B \in \mathcal{M}_{m\times n}, C \in \mathcal{M}_{n\times p}$

$$(A + B) \cdot C = A \cdot C + B \cdot C.$$

4. $\forall A \in \mathcal{M}_{m\times n}, B, C \in \mathcal{M}_{n\times p}$

$$A \cdot (B + C) = A \cdot B + A \cdot C.$$

5. $\forall A \in \mathcal{M}_{m\times n}$

$$\begin{array}{lcl} I_m \cdot A = A & , & 0_{p\times m} \cdot A_{m\times n} = 0_{p\times n} \\ A \cdot I_n = A & , & A_{m\times n} \cdot 0_{n\times p} = 0_{m\times p} \end{array}.$$

6. $\forall A \in \mathcal{M}_{m\times n}, B \in \mathcal{M}_{n\times p}$

$$(A \cdot B)^t = B^t \cdot A^t.$$

7. $\forall A, B \in \mathcal{M}_n, A \cdot B \in \mathcal{M}_n.$

8. *Given* $A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}_{n\times n}$, $B = \begin{pmatrix} b_1 & & & \\ & b_2 & & \\ & & \ddots & \\ & & & b_n \end{pmatrix}_{n\times n}$, *diagonal matrices of* $\mathcal{M}_n$, *we have that:*

$$A \cdot B = \begin{pmatrix} a_1 \cdot b_1 & & & \\ & a_2 \cdot b_2 & & \\ & & \ddots & \\ & & & a_n \cdot b_n \end{pmatrix}_{n\times n}.$$

Example 111. If we consider the matrices

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix},$$

we can perform their product in the usual way. However, since both are diagonal, it is sufficient to multiply the elements of the main diagonal so that

$$A \cdot B = \begin{pmatrix} 3\cdot 2 & 0 & 0 & 0 \\ 0 & (-1)\cdot(-3) & 0 & 0 \\ 0 & 0 & 4\cdot 1 & 0 \\ 0 & 0 & 0 & 2\cdot 5 \end{pmatrix} = \begin{pmatrix} 6 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix}.$$

Remark. Given two matrices $A \in \mathcal{M}_{m\times n}, B \in \mathcal{M}_{p\times r}$, we may ask when it will be possible to perform both the operation $A \cdot B$ and the operation $B \cdot A$. We have that:

$$\left.\begin{array}{lll} \text{if it is possible to perform } A_{m\times n} \cdot B_{p\times r} & \Rightarrow & n = p \\ \text{if it is possible to perform } B_{p\times r} \cdot A_{m\times n} & \Rightarrow & r = m \end{array}\right\} \Rightarrow \left\{\begin{array}{l} A \in \mathcal{M}_{m\times n} \\ B \in \mathcal{M}_{n\times m} \end{array}\right.$$

$$\text{in which case } \left\{\begin{array}{l} A \cdot B \in \mathcal{M}_{m\times m} \\ B \cdot A \in \mathcal{M}_{n\times n} \end{array}\right. .$$

In particular, if $A, B \in \mathcal{M}_n$ (are of type $n \times n$) we can calculate both $A \cdot B$ and $B \cdot A$ and furthermore we obtain $A \cdot B, B \cdot A \in \mathcal{M}_n$.

In general the product of matrices does not satisfy the commutative property ($A \cdot B = B \cdot A$) since we will not always be able to calculate $A \cdot B$ and $B \cdot A$ and even when it is possible we will not always obtain the same result.

Examples 112.

1) $\begin{pmatrix} 3 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix}_{2\times 3} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}_{3\times 1} = \begin{pmatrix} 7 \\ 4 \end{pmatrix}$.

$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}_{3\times 1} \cdot \begin{pmatrix} 3 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix}_{2\times 3}$ cannot be calculated.

2) If $A = \begin{pmatrix} 3 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}$ since they are of type 2×3 and 3×2, we can calculate both $A \cdot B$ and $B \cdot A$. However we have that:

$\star\ A \cdot B = \begin{pmatrix} 3 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 7 & 3 \\ 4 & -1 \end{pmatrix}.$

$\star\ B \cdot A = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 2 \\ 4 & 1 & 4 \\ 7 & 3 & 2 \end{pmatrix}.$

We thus obtain that $A \cdot B \neq B \cdot A$ and furthermore they are not even matrices of the same type.

3) If $A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$ then we can calculate both $A \cdot B$ and $B \cdot A$ and furthermore $A \cdot B, B \cdot A \in \mathcal{M}_2$, however:

$\star\ A \cdot B = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ -1 & -1 \end{pmatrix}.$

$\star\ B \cdot A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$

So in this case $A \cdot B$ and $B \cdot A$ are of the same type but are different matrices.

4) Given any matrix $A \in \mathcal{M}_n$ it will be verified that

$$A \cdot I_n = I_n \cdot A = A$$

so in this situation the products $A \cdot I_n$ and $I_n \cdot A$ do coincide.

Power of matrices

Given a matrix $A \in \mathcal{M}_{m\times n}$ it will be possible to perform the product

$$A_{m\times n} \cdot A_{m\times n}$$

only when $m = n$ in which case A would be a square matrix of order n. In those cases where A is not a square matrix it will never be possible to perform the operation $A \cdot A$.

We see then that it is only possible to calculate the product of a matrix by itself when it is square and then, when performing this product, we obtain again a square matrix of the same type as the initial one. This makes it possible for us to multiply a square matrix by itself as many times as we want. This leads to the concept of power of a square matrix that we introduce in the next definition and which, as we will see later, constitutes one of the most important tools in the construction of matrix models.

Definition 113. Given $A \in \mathcal{M}_n$ we define, for $k \in \mathbb{N}$,

$$A^k = A \cdot A \cdot A \cdot \overset{k)}{\cdots} \cdot A \in \mathcal{M}_n.$$

Examples 114.
1) The matrix

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \\ 2 & -1 & 1 \end{pmatrix}$$

is square of order three. Being square we can calculate its powers. For example,

$$A^2 = A \cdot A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \\ 2 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \\ 2 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 1 & 2 \\ 5 & 0 & 0 \\ 5 & 2 & -1 \end{pmatrix}.$$

We can also calculate A^3. To do this we will take advantage of the calculation done with A^2 in the following way,

$$A^3 = A \cdot A \cdot A = A \cdot A^2 = \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \\ 2 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 5 & 1 & 2 \\ 5 & 0 & 0 \\ 5 & 2 & -1 \end{pmatrix} = \begin{pmatrix} 15 & 2 & 4 \\ 10 & 5 & 0 \\ 10 & 4 & 3 \end{pmatrix}.$$

In reality, to calculate the successive powers of the matrix A (A^4, A^5, etc.), we can repeat this process by multiplying the last power obtained again by the matrix A to thus obtain the next power. For example, if we repeat the process once more we obtain A^4:

$$A^4 = A \cdot A^3 = \underbrace{\begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 2 \\ 2 & -1 & 1 \end{pmatrix}}_{\text{The matrix } A,} \cdot \underbrace{\begin{pmatrix} 15 & 2 & 4 \\ 10 & 5 & 0 \\ 10 & 4 & 3 \end{pmatrix}}_{\substack{\text{multiplied by the} \\ \text{previous power, } A^3,}} = \underbrace{\begin{pmatrix} 40 & 9 & 8 \\ 25 & 5 & 10 \\ 30 & 3 & 11 \end{pmatrix}}_{\substack{\text{provides the next} \\ \text{power, } A^4.}}.$$

2) The matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 2 & -2 \end{pmatrix}$$

is square of order 3 and we can calculate the product of it by itself and any of its powers. Let's calculate the powers A^2, A^3 and A^4. To do this we will start with A^2. We have that

$$A^2 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 2 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 2 & -2 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 3 \\ 2 & 5 & -6 \\ -1 & -4 & 9 \end{pmatrix}.$$

Once we have calculated A^2 we can multiply once more by A to obtain A^3 in the form

$$A^3 = A^{2+1} = A^2 \cdot A = \begin{pmatrix} 2 & 2 & 3 \\ 2 & 5 & -6 \\ -1 & -4 & 9 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 2 & -2 \end{pmatrix} = \begin{pmatrix} 5 & 8 & 0 \\ -4 & -13 & 24 \\ 8 & 20 & -27 \end{pmatrix}.$$

Finally, applying the same procedure once more,

$$A^4 = A^{3+1} = A^3 \cdot A = \begin{pmatrix} 5 & 8 & 0 \\ -4 & -13 & 24 \\ 8 & 20 & -27 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 2 & -2 \end{pmatrix} = \begin{pmatrix} 5 & 2 & 21 \\ 20 & 53 & -78 \\ -19 & -58 & 102 \end{pmatrix}.$$

On the other hand, it is important to highlight the fact that the power of matrices cannot be solved by raising the elements of the matrix one by one. That is, in general, we cannot calculate the power A^k by

$$\begin{pmatrix} 1^k & 2^k & 1^k \\ 0^k & (-1)^k & 2^k \\ 1^k & 2^k & (-2)^k \end{pmatrix}.$$

To realize this, it is enough to check this for the power A^4 that we calculated before:

$$A^4 = \underbrace{\begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 1 & 2 & -2 \end{pmatrix}^4}_{= \begin{pmatrix} 5 & 2 & 21 \\ 20 & 53 & -78 \\ -19 & -58 & 102 \end{pmatrix}} \neq \underbrace{\begin{pmatrix} 1^4 & 2^4 & 1^4 \\ 0^4 & (-1)^4 & 2^4 \\ 1^4 & 2^4 & (-2)^4 \end{pmatrix}}_{= \begin{pmatrix} 1 & 16 & 1 \\ 0 & 1 & 16 \\ 1 & 16 & 16 \end{pmatrix}}.$$

In the previous example the basic procedure for calculating powers of a matrix is described. It can be observed that the calculation of these powers leads to the performance of numerous matrix products and therefore involves a high number of operations. Even if we have to calculate a power with a moderately high exponent, we will find that it is impossible to perform the calculation by hand. This makes the calculation of matrix powers a difficult operation that requires the use of other more sophisticated techniques that we will study in the chapter on diagonalization.

In the following property, among other things, we will see that there is an exception to what was said in the previous paragraph. When the matrix for which we want to calculate the power is diagonal, then the calculation is simplified notably.

Properties 115.

1. *Given $A \in \mathcal{M}_n$ and $k, p \in \mathbb{N}$*
$$A^k \cdot A^p = A^p \cdot A^k = A^{k+p}.$$

2. *Given the diagonal matrix* $A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}_{n\times n} \in \mathcal{M}_n$ *and $k \in \mathbb{N}$ it is verified that:*
$$A^k = \begin{pmatrix} a_1^k & & & \\ & a_2^k & & \\ & & \ddots & \\ & & & a_n^k \end{pmatrix}_{n\times n}.$$

Examples 116.

1) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}^4 = \begin{pmatrix} 2^4 & 0 & 0 \\ 0 & 3^4 & 0 \\ 0 & 0 & (-1)^4 \end{pmatrix} = \begin{pmatrix} 16 & 0 & 0 \\ 0 & 81 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$

2) Given $A \in \mathcal{M}_n$ it is verified that

$$\begin{aligned} A^3 \cdot A^2 &= (A \cdot A \cdot A) \cdot (A \cdot A) = A \cdot A \cdot A \cdot A \cdot A = A^5, \\ A^2 \cdot A^3 &= (A \cdot A) \cdot (A \cdot A \cdot A) = A \cdot A \cdot A \cdot A \cdot A = A^5. \end{aligned}$$

Therefore $A^3 \cdot A^2 = A^2 \cdot A^3 = A^5$ as indicated by the first of the properties stated above. This example gives us an idea of how this property can be demonstrated in its most general form.

3) Point 2 of the previous property is applicable only to diagonal matrices. For arbitrary matrices such property, in general, will not hold. Thus for example:

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 1 & 2 & 1 \end{pmatrix}^3 = \begin{pmatrix} 3 & 6 & 2 \\ 1 & -1 & 3 \\ 5 & 8 & 5 \end{pmatrix} \neq \begin{pmatrix} 1^3 & 2^3 & 0^3 \\ 0^3 & (-1)^3 & 1^3 \\ 1^3 & 2^3 & 1^3 \end{pmatrix} = \begin{pmatrix} 1 & 8 & 0 \\ 0 & -1 & 1 \\ 1 & 8 & 1 \end{pmatrix}.$$

4.2.4 Inverse matrix

Considering two matrices A and B of the appropriate types, we can calculate their sum, $A+B$, their difference, $A-B$, or their product, $A \cdot B$. However, we have not yet defined any operation that we can call 'division of matrices'. This is what we are going to attempt in this section by defining the concept of inverse of a matrix. Due to the type of operations involved, only a square matrix can have an inverse. To introduce these new concepts we will begin by making some considerations:

- Given any matrix $A \in \mathcal{M}_n$ we know that
$$A \cdot I_n = I_n \cdot A = A,$$
a property analogous to that verified by the number 1 within the real numbers ($\forall r \in \mathbb{R},\ 1 \cdot r = r \cdot 1 = r$). We can then admit that the matrix I_n plays the role of 1 within the set of square matrices of order n. That is, I_n is the unit of $\mathcal{M}_n$.

- Given $a, b \in \mathbb{R},\ b \neq 0$, the division of a by b can be calculated as
$$\frac{a}{b} = a \cdot \frac{1}{b} = a \cdot b^{-1},$$
where b^{-1} is what is called the inverse of the number b. Therefore, when dividing two numbers, what really interests us is to calculate the inverse of one of them.

- Given $b \in \mathbb{R},\ b \neq 0$, we know that its inverse is another real number that we write as b^{-1} and that is the only number that satisfies
$$b \cdot b^{-1} = b^{-1} \cdot b = 1.$$
The inverse of b is that number by which b must be multiplied to obtain 1.

- Not every real number has an inverse since it is not possible to calculate $0^{-1} = \frac{1}{0}$ because there is no number x such that
$$0 \cdot x = 1.$$

Due to all the above, it seems clear that, given a matrix $A \in \mathcal{M}_n$, if we want to define A^{-1}, we will have to find another matrix, $B \in \mathcal{M}_n$, such that
$$A \cdot B = B \cdot A = I_n$$
and then that matrix B will be the inverse of A, that is, $A^{-1} = B$.

Definition 117. Given $A \in \mathcal{M}_n$, if it exists, we call the inverse matrix of A and denote it by A^{-1} the unique matrix that satisfies:
$$A^{-1} \cdot A = A \cdot A^{-1} = I_n.$$

Examples 118.

1) Given $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ consider the matrix $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ and then we have that

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2, \\ \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2. \end{aligned}$$

Therefore, according to the definition we have given, we have that:
$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

2) Given $\begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ if we take $\begin{pmatrix} -1 & 2 & 2 \\ 1 & -1 & -1 \\ 1 & -1 & -2 \end{pmatrix}$ we obtain:

$$\begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 & 2 \\ 1 & -1 & -1 \\ 1 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3,$$
$$\begin{pmatrix} -1 & 2 & 2 \\ 1 & -1 & -1 \\ 1 & -1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3$$

and therefore $\begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 2 & 2 \\ 1 & -1 & -1 \\ 1 & -1 & -2 \end{pmatrix}$.

3) Given $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -5 \end{pmatrix}$, taking into account properties we know about the product of diagonal matrices, it is easy to calculate the inverse by taking the inverses of the elements of the main diagonal of B:

$$\begin{pmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -5 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{9} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{-5} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3,$$
$$\begin{pmatrix} \frac{1}{9} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{-5} \end{pmatrix} \cdot \begin{pmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3$$

and therefore $B^{-1} = \begin{pmatrix} \frac{1}{9} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{-5} \end{pmatrix}$.

Note that this technique cannot be used with matrices that are not diagonal as evidenced in parts **1)** and **2)** of **Examples 118** in which the inverse matrix is not the matrix formed by the inverses of the elements of the initial matrix.

If we think about what would have happened if some of the elements of the main diagonal of B had been zero, it also seems clear, in view of how we obtained B^{-1} before, that in such a case it would not be possible to calculate the inverse.

4) Let us determine whether $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ has an inverse matrix. If A has an inverse, it will also be a square matrix of order 2 and therefore of the form

$$A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then we will have that

$$A \cdot A^{-1} = I_2 \Rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = I_2 \Rightarrow \begin{pmatrix} a+c & b+d \\ a+c & b+d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \left\{ \begin{array}{ll} a+c=1 & b+d=1 \\ a+c=0 & b+d=0 \end{array} \right. ,$$

which is impossible since it is evident that the quantity $a+c$ cannot be simultaneously equal to 1 and equal to 0. We therefore deduce that the matrix A does not have an inverse matrix.

5) Let's try to calculate the inverse of the matrix $A = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$.

To do this we will apply the same technique that we used in the previous example. In this way, if A has an inverse we know that it will also be a square matrix of order 2 and therefore must be of the form

$$A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Since the inverse matrix must satisfy $A \cdot A^{-1} = I_2$ we will have that

$$\begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 3a+c & 3b+d \\ 2a+c & 2b+d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow\Rightarrow \begin{cases} 3a+c=1, & 3b+d=0, \\ 2a+c=0, & 2b+d=1. \end{cases}$$

$$\overset{\text{Solving the system}}{\Rightarrow} \begin{cases} a=1, & b=-1, \\ c=-2, & d=3. \end{cases}$$

See that in the end we obtain a linear system with four equations and four unknowns that is easily solved so that finally we have calculated the inverse matrix which will be

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

Definition 119. We say that $A \in \mathcal{M}_n$ is a regular matrix if the inverse matrix of A exists and otherwise we say that A is a singular or non-regular matrix.

Example 120. $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ is a singular while $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ or $\begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$ are regular.

In parts **4)** and **5** of **Examples 118** we have presented the basic technique to decide whether a matrix is regular or not and, if it is, to calculate its inverse. As we see, if we apply this technique, the calculation of the inverse depends on the resolution of a system that in the case of a 2×2 matrix will have four equations and four unknowns. The problem posed by this technique lies in the fact that, for matrices of higher order, the number of equations and variables that will appear in the system multiplies considerably. Thus, for a matrix of order 3 we would have 9 equations and for one of order 4 it would be 16. For this reason this method is valid only in simple cases such as that of order 2 matrices. Once, in the following sections of this topic, we introduce the techniques for handling matrices through elementary operations, we will have a more effective method for calculating inverse matrices.

Let's see some important properties of the inverse matrix.

Properties 121.

i) If $A, B \in \mathcal{M}_n$ are regular then the matrix $A \cdot B$ is also regular and furthermore it is verified that:

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}.$$

ii) If $A \in \mathcal{M}_n$ is regular then A^t is also regular and furthermore it is verified that:

$$(A^t)^{-1} = (A^{-1})^t.$$

iii) Given the diagonal matrix $A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}$ *such that* $a_1 \neq 0$, $a_2 \neq 0, \ldots, a_n \neq 0$, *we have that* *A is regular and furthermore:*

$$A^{-1} = \begin{pmatrix} \frac{1}{a_1} & & & \\ & \frac{1}{a_2} & & \\ & & \ddots & \\ & & & \frac{1}{a_n} \end{pmatrix}.$$

iv) If $A \in \mathcal{M}_n$ *is regular, then* A^{-1} *is also regular and*

$$(A^{-1})^{-1} = A.$$

v) If $A \in \mathcal{M}_n$ *is regular and we take* $r \in \mathbb{R}$, $r \neq 0$, *then* $r \cdot A$ *is regular and it is verified that:*

$$(r \cdot A)^{-1} = \frac{1}{r} \cdot A^{-1}.$$

Examples 122. In the following examples we will resort to the different points of **Properties 121**:
1) Let's calculate the inverse of some matrices taking advantage of some previous operations and the properties already seen. We have that

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

We know through example 118 on page 134 the following inverse

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Also,

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^t,$$

so, using property *ii)*,

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^t\right)^{-1} = \left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1}\right)^t = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}^t = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

and finally we can also use property *i)* to calculate

$$\begin{aligned} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1} &= \left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right)^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}. \end{aligned}$$

2) Given $I_n \in \mathcal{M}_n$ we know that $I_n \cdot I_n = I_n$ so the identity matrix is regular and its inverse is itself:

$$(I_n)^{-1} = I_n.$$

On the other hand, since I_n is a diagonal matrix, we can also use property *iii)* to reach the same conclusion about its inverse.

3) Given $A, B, C \in \mathcal{M}_n$ we are going to calculate the inverse $(A \cdot B \cdot C)^{-1}$. To do this we will repeatedly use property *i)*.

$$\begin{aligned}(A \cdot B \cdot C)^{-1} &= ((A \cdot B) \cdot C)^{-1} = (\text{property } i)) = C^{-1} \cdot (A \cdot B)^{-1} \\ &= (\text{property } i)) = C^{-1} \cdot (B^{-1} \cdot A^{-1}) = C^{-1} \cdot B^{-1} \cdot A^{-1}.\end{aligned}$$

In general reiterating this process it is easy to verify that given any number, k, of regular square matrices, $A_1, A_2, \ldots, A_k \in \mathcal{M}_n$, it is verified that

$$(A_1 \cdot A_2 \cdot \cdots \cdot A_k)^{-1} = A_k^{-1} \cdot \cdots \cdot A_2^{-1} \cdot A_1^{-1}.$$

4.3 Linear combinations, linear independence. Rank of a matrix

The analysis of a phenomenon will generally be based on the data related to that phenomenon that we obtain by measuring the variables involved in it. Determining which variables are essential or not, which ones can be obtained or explained through others, and in general measuring the amount of information we possess about the phenomenon are essential tasks that are based on the concepts of linear combination, dependence and independence that we will see in this section. We will begin by showing an example that highlights all these issues.

Example 123. Suppose we are studying the vehicle fleet in different cities with the aim of making decisions about the possibility of installing new companies in this sector. We conduct the study in seven cities that we will call A, B, C, D, E, F and G. In each of them we will initially study two variables:

$$\begin{aligned} N_C = &\text{ Number of cars present in the city,} \\ N_M = &\text{ Number of motorcycles.}\end{aligned}$$

After the corresponding data collection we obtain the following values for these variables in each city (expressed in thousands of vehicles of each type):

	N_C	N_M
City A	7	6
City B	8	5
City C	10	5
City D	6	6
City E	4	5
City F	20	10
City G	9	5

We intend to analyze the recycling of tires and motor-derived waste. For this it is reasonable that we study in each city two new variables:

$$\begin{aligned} N_R = &\text{ Number of tires circulating in the city's vehicles,} \\ N_m = &\text{ Number of engines in use.}\end{aligned}$$

We could perform new data collections in the cities of the study to obtain the information of these other two variables, however, it is evident that each car has four wheels and each motorcycle two, and that there will be a single engine per vehicle. Therefore, in this case, it is not necessary to take more data since, knowing the variables corresponding to the number of cars and motorcycles, N_C and N_M, we can calculate the other two variables because clearly we will have that

$$N_R = 4N_C + 2N_M \qquad \text{and} \qquad N_m = N_C + N_M. \tag{4.2}$$

In this way if we obtained the data for all variables we would get

	N_C	N_M	N_R	N_m
City A	7	6	40	13
City B	8	5	42	13
City C	10	5	50	15
City D	6	6	36	12
City E	4	5	26	9
City F	20	10	100	30
City G	9	5	46	14

and it is possible to verify how for each city the relations given in (4.2) are satisfied. Actually, each of the four variables is a 7-tuple since it contains seven data, one per city, and using the operations we know for tuples we have that

$$\underbrace{\begin{pmatrix} 40 \\ 42 \\ 50 \\ 36 \\ 26 \\ 100 \\ 46 \end{pmatrix}}_{=N_R} = 4 \underbrace{\begin{pmatrix} 7 \\ 8 \\ 10 \\ 6 \\ 4 \\ 20 \\ 9 \end{pmatrix}}_{=N_C} +2 \underbrace{\begin{pmatrix} 6 \\ 5 \\ 5 \\ 6 \\ 5 \\ 10 \\ 5 \end{pmatrix}}_{=N_M} \qquad \text{and} \qquad \underbrace{\begin{pmatrix} 13 \\ 13 \\ 15 \\ 12 \\ 9 \\ 30 \\ 14 \end{pmatrix}}_{=N_m} = \underbrace{\begin{pmatrix} 7 \\ 8 \\ 10 \\ 6 \\ 4 \\ 20 \\ 9 \end{pmatrix}}_{=N_C} + \underbrace{\begin{pmatrix} 6 \\ 5 \\ 5 \\ 6 \\ 5 \\ 10 \\ 5 \end{pmatrix}}_{=N_M}.$$

We can verify that the information of N_R and N_m depends on N_C and N_M and therefore it is not necessary for us to take data in each city for these variables, we simply have to obtain the information for N_R and N_m by combining what we already have in N_C and N_M through the dependence relations given by the equations in (4.2). Actually, we can combine N_C and N_M to obtain many other variables but all of them will contain superfluous information that can be calculated from the data we already have. Other variables that are obtained by combination of N_C and N_M could be:

- N_P = Maximum number of transportable passengers,
- N_F = Number of nighttime illumination headlights.

Under the assumption that each car can transport five passengers and has two headlights and that each motorcycle transports a maximum of two with a single headlight, the following dependence relations are evident,

$$N_P = 5N_C + 2N_M \qquad \text{and} \qquad N_F = 2N_C + N_M,$$

through which we can calculate the values taken by N_P and N_F in the different cities by tuple calculus in the following way

$$\underbrace{\begin{pmatrix} 47 \\ 50 \\ 60 \\ 42 \\ 30 \\ 120 \\ 55 \end{pmatrix}}_{=N_P} = 5 \underbrace{\begin{pmatrix} 7 \\ 8 \\ 10 \\ 6 \\ 4 \\ 20 \\ 9 \end{pmatrix}}_{=N_C} +2 \underbrace{\begin{pmatrix} 6 \\ 5 \\ 5 \\ 6 \\ 5 \\ 10 \\ 5 \end{pmatrix}}_{=N_M} \qquad \text{and} \qquad \underbrace{\begin{pmatrix} 20 \\ 21 \\ 25 \\ 18 \\ 13 \\ 50 \\ 23 \end{pmatrix}}_{=N_F} = 2 \underbrace{\begin{pmatrix} 7 \\ 8 \\ 10 \\ 6 \\ 4 \\ 20 \\ 9 \end{pmatrix}}_{=N_C} + \underbrace{\begin{pmatrix} 6 \\ 5 \\ 5 \\ 6 \\ 5 \\ 10 \\ 5 \end{pmatrix}}_{=N_M}.$$

Actually any variable, N, that is obtained as a combination of N_C and N_M will be of the form

$$N = \alpha N_C + \beta N_M$$

and its values can be calculated by the tuple operation

$$N = \alpha \underbrace{\begin{pmatrix} 7 \\ 8 \\ 10 \\ 6 \\ 4 \\ 20 \\ 9 \end{pmatrix}}_{=N_C} + \beta \underbrace{\begin{pmatrix} 6 \\ 5 \\ 5 \\ 6 \\ 5 \\ 10 \\ 5 \end{pmatrix}}_{=N_M}.$$

For different values of α and β we can obtain an infinity of variables that are combinations of N_C and N_M but in all cases their information will be superfluous once we know these last two.

On the other hand, it seems clear that the number of cars in a city is completely independent of the number of motorcycles. That is, knowing the number of cars it is impossible to calculate the number of motorcycles and equally in the opposite case. Thus, the variables N_C and N_M are independent of each other and both are essential so we need to take the data of both without one being obtainable from the other. That is, there is no formula of the type

$$N_C = \alpha N_M \qquad \text{or} \qquad N_M = \alpha N_C$$

that allows obtaining N_C as a combination of N_M or vice versa.

Thus, it seems clear that in this problem the essential variables are N_C and N_M from which we can derive others as combinations.

The previous example serves as an introduction to the concepts of combination, linear dependence and independence that we introduce in the following definition. It is evident that instead of studying seven cities, the study could be extended to eight, nine or, in general, n. In such a case, each variable would be an n-tuple with n data, one for each city. Likewise, more than two variables can participate in our study, instead we could have three, four or, in general, m which would give us m, n-tuples, one for each variable.

Definition 124. Consider the n-tuples $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$. Then:

i) We say that the n-tuple, $w \in \mathbb{R}^n$, is a linear combination of $v_1, v_2, \ldots, v_m$ if

$$w = \alpha_1 v_1 + \alpha_2 v2 + \cdots + \alpha_m v_m$$

for certain real numbers $\alpha_1, \alpha_2, \ldots, \alpha_m$ that we call coefficients of the combination. The set of all linear combinations of $v_1, v_2, \ldots, v_m$ is denoted $\langle v_1, v_2, \ldots, v_m \rangle$.

ii) We say that $v_1, v_2, \ldots, v_m$ are linearly independent if none of them can be obtained as a linear combination of the others. We say that a single tuple is independent (i.e., $m = 1$) provided it is not null.

iii) We say that $v_1, v_2, \ldots, v_m$ are linearly dependent if they are not independent.

These concepts correspond directly with the ideas we have presented in the example at the beginning of the section. Indeed, applying this new terminology in the previous example we can say that:

- The tuples N_R, N_m, N_P and N_F are linear combinations of N_C and N_M. In other words,

  $$N_R, N_m, N_P, N_F \in \langle N_C, N_M \rangle.$$

  We can also obtain many other combinations of N_C and N_M and all of them will be of the form $\alpha N_C + \beta N_M$ for certain numbers $\alpha, \beta \in \mathbb{R}$. Therefore,

  $$\langle N_C, N_M \rangle = \{\alpha N_C + \beta N_M : \alpha, \beta \in \mathbb{R}\}.$$

- The set formed by the tuples $N_C, N_M, N_R, N_m, N_P, N_F$ is linearly dependent since some of them (the last four) can be obtained as linear combinations of the others (of the first two). Consequently this set we know that there are tuples (variables) that are superfluous whose information can be obtained from the others.
- The tuples N_C and N_M are independent since neither is a linear combination of the other. Consequently the information contained in these two tuples is essential and neither of them can be considered superfluous.

From several tuples we can generate many others through linear combinations. In the previous example, we have seen that from N_C and N_M we can obtain many other new tuples such as N_R, N_m, N_P or N_F and actually we could obtain an infinity since simply by modifying the coefficients we use in each linear combination we will obtain a new one. That is why the set $\langle N_C, N_M\rangle$ has infinitely many elements. The same happens when we have any other tuples, by combining them we can obtain an unlimited amount of new tuples and the set of their linear combinations will also be infinite.

The following property is essential since it emphasizes the idea that a tuple that can be obtained as a linear combination of others is superfluous in a certain sense. Specifically we will see that, when generating linear combinations, one of such tuples is unnecessary and can be eliminated.

Property 125. *Given the n-tuples w and $v_1, v_2, \ldots, v_m$, it holds that*

$$\underbrace{w \in \langle v_1, v_2, \ldots, v_m\rangle}_{\substack{\text{If } w \text{ is obtained as} \\ \text{combination of } v_1,v_2,\ldots,v_m,}} \Leftrightarrow \underbrace{\langle w, v_1, v_2, \ldots, v_m\rangle}_{\substack{\text{all combinations} \\ \text{obtained using } w,}} = \underbrace{\langle v_1, v_2, \ldots, v_m\rangle.}_{\substack{\text{can also be obtained} \\ \text{if we remove } w}}\ .$$

Therefore, when we have a set of tuples and we find out that they are dependent, we will also know that, when obtaining tuples through linear combinations, said set can be simplified since we will have superfluous tuples. For example, combining N_C, N_M and N_R we can obtain many tuples that will form the set

$$\langle N_C, N_M, N_R\rangle.$$

Now, we know that these three tuples, $\{N_C, N_M, N_R\}$ form a dependent set since N_R can be obtained as a combination of the other two. Consequently, N_R is superfluous and we can eliminate it since

$$\underbrace{\langle N_C, N_M, N_R\rangle}_{\substack{\text{all combinations} \\ \text{obtained using } N_R,}} = \underbrace{\langle N_C, N_M\rangle.}_{\substack{\text{can also be obtained} \\ \text{if we remove } N_R}}\ .$$

In other words, being dependent, in the initial set of combinations we can eliminate N_R,

$$\langle N_C, N_M, \not{N_R}\rangle.$$

Let's see some additional examples.

Examples 126.

1) Consider the columns

$$\begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}.$$

If we add them multiplied by the numbers 5, 2 and -1 we obtain

$$5 \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} + 2 \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + (-1) \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 15 \\ 10 \\ -5 \end{pmatrix} + \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ -3 \end{pmatrix} = \boxed{\begin{pmatrix} 18 \\ 11 \\ -8 \end{pmatrix}}.$$

Therefore, the tuple $(18, 11, -8)$ is obtained as a combination of the three initial ones with the coefficients of that combination being 5, 2 and -1. Writing this in another way, we can put that

$$\begin{pmatrix}18\\11\\-8\end{pmatrix} \in \langle \begin{pmatrix}3\\2\\-1\end{pmatrix}, \begin{pmatrix}2\\0\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\3\end{pmatrix} \rangle.$$

2) Let us now take the 5-tuples

$$\begin{pmatrix}2 & 3 & 0 & 0 & 6\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}-1 & 2 & 3 & 0 & 1\end{pmatrix}.$$

Again we can combine them to obtain a row different from the initial ones. For example we can multiply by 3 and 2,

$$\boxed{3}\begin{pmatrix}2 & 3 & 0 & 0 & 6\end{pmatrix} + \boxed{2}\begin{pmatrix}-1 & 2 & 3 & 0 & 1\end{pmatrix} =$$

$$= \begin{pmatrix}6 & 9 & 0 & 0 & 18\end{pmatrix} + \begin{pmatrix}-2 & 4 & 6 & 0 & 2\end{pmatrix} = \boxed{\begin{pmatrix}4 & 13 & 6 & 0 & 20\end{pmatrix}}.$$

If we had chosen different coefficients we would have obtained as a result a row also different. In this way, if we now take 4 and -1 we have

$$\boxed{4}\begin{pmatrix}2 & 3 & 0 & 0 & 6\end{pmatrix} + \boxed{(-1)}\begin{pmatrix}-1 & 2 & 3 & 0 & 1\end{pmatrix} =$$

$$= \begin{pmatrix}8 & 12 & 0 & 0 & 24\end{pmatrix} + \begin{pmatrix}1 & -2 & -3 & 0 & -1\end{pmatrix} = \boxed{\begin{pmatrix}9 & 10 & -3 & 0 & 23\end{pmatrix}}.$$

In short we have that

$$\begin{pmatrix}4 & 13 & 6 & 0 & 20\end{pmatrix}, \begin{pmatrix}9 & 10 & -3 & 0 & 23\end{pmatrix} \in \left\langle \begin{pmatrix}2 & 3 & 0 & 0 & 6\end{pmatrix}, \begin{pmatrix}-1 & 2 & 3 & 0 & 1\end{pmatrix} \right\rangle.$$

Actually, by changing the coefficients we use to combine the two initial rows we could generate an infinite amount of new linear combinations and therefore we know that the set

$$\left\langle \begin{pmatrix}2 & 3 & 0 & 0 & 6\end{pmatrix}, \begin{pmatrix}-1 & 2 & 3 & 0 & 1\end{pmatrix} \right\rangle$$

has infinitely many elements. Despite this, it is worth asking whether it is possible to obtain any 5-tuple or if on the contrary there exist tuples of $\mathbb{R}^5$ that cannot be obtained by combining the two initial ones. In this case it is easy to verify that the second option occurs since we can directly verify that through the two rows we have it is not possible to obtain the 5-tuple $\begin{pmatrix}0 & 0 & 0 & 1 & 0\end{pmatrix}$ as a linear combination since, no matter what numbers a_1 and a_2 we choose to perform the combination, we will always have

$$\underbrace{\boxed{a_1}\begin{pmatrix}2 & 3 & 0 & 0 & 6\end{pmatrix} + \boxed{a_2}\begin{pmatrix}-1 & 2 & 3 & 0 & 1\end{pmatrix}}_{= \begin{pmatrix}2a_1 - a_2 & 3a_1 + 2a_2 & 3a_2 & \boxed{0} & 6a_1 + a_2\end{pmatrix}} \quad \text{and} \quad \underbrace{\begin{pmatrix}0 & 0 & 0 & \boxed{1} & 0\end{pmatrix}}_{}$$

the result always has a 0 in the fourth position — the row we intend to obtain has a 1 in the fourth position

and therefore no combination of the two initial rows will be able to produce the row in question. That is,

$$\begin{pmatrix}0 & 0 & 0 & 1 & 0\end{pmatrix} \notin \left\langle \begin{pmatrix}2 & 3 & 0 & 0 & 6\end{pmatrix}, \begin{pmatrix}-1 & 2 & 3 & 0 & 1\end{pmatrix} \right\rangle$$

and consequently

$$\left\langle \begin{pmatrix}2 & 3 & 0 & 0 & 6\end{pmatrix}, \begin{pmatrix}-1 & 2 & 3 & 0 & 1\end{pmatrix} \right\rangle \neq \mathbb{R}^5.$$

That is, not every 5-tuple can be obtained as a linear combination of $\begin{pmatrix}2 & 3 & 0 & 0 & 6\end{pmatrix}$ and $\begin{pmatrix}-1 & 2 & 3 & 0 & 1\end{pmatrix}$.

3) Consider the columns $\begin{pmatrix} 2\\0\\-1\\1 \end{pmatrix}$, $\begin{pmatrix} 1\\2\\1\\1 \end{pmatrix}$, $\begin{pmatrix} 5\\-1\\-7\\-3 \end{pmatrix}$ and $\begin{pmatrix} 0\\1\\2\\2 \end{pmatrix}$. By performing linear combinations of these columns it is possible to obtain new ones. The set of all their linear combinations will be

$$\langle \begin{pmatrix} 2\\0\\-1\\1 \end{pmatrix}, \begin{pmatrix} 1\\2\\1\\1 \end{pmatrix}, \begin{pmatrix} 5\\-1\\-7\\-3 \end{pmatrix}, \begin{pmatrix} 0\\1\\2\\2 \end{pmatrix} \rangle =$$
$$= \{a_1 \begin{pmatrix} 2\\0\\-1\\1 \end{pmatrix} + a_2 \begin{pmatrix} 1\\2\\1\\1 \end{pmatrix} + a_3 \begin{pmatrix} 5\\-1\\-7\\-3 \end{pmatrix} + a_4 \begin{pmatrix} 0\\1\\2\\2 \end{pmatrix} : a_1, a_2, a_3, a_4 \in \mathbb{R}\}.$$

For example, taking $a_1 = 2$, $a_2 = -1$, $a_3 = 1$ and $a_4 = 3$ we obtain the linear combination

$$2 \begin{pmatrix} 2\\0\\-1\\1 \end{pmatrix} - 1 \begin{pmatrix} 1\\2\\1\\1 \end{pmatrix} + 2 \begin{pmatrix} 5\\-1\\-7\\-3 \end{pmatrix} + 3 \begin{pmatrix} 0\\1\\2\\2 \end{pmatrix} = \begin{pmatrix} 13\\-1\\-11\\1 \end{pmatrix}.$$

The question is whether it is possible to obtain the same linear combinations with fewer columns or, said in another way, if there is any of the four columns that is superfluous. Answering this question turns out to be simple if we realize that the third column can be obtained by combining the other three in the form

$$\begin{pmatrix} 5\\-1\\-7\\-3 \end{pmatrix} = 2 \begin{pmatrix} 2\\0\\-1\\1 \end{pmatrix} + 1 \begin{pmatrix} 1\\2\\1\\1 \end{pmatrix} - 3 \begin{pmatrix} 0\\1\\2\\2 \end{pmatrix}. \tag{4.3}$$

Directly this means that the four initial columns are dependent since one of them (the third) is obtained by combining the others and in such a case, if we apply **Property 125**, we know then that we can dispense with that third column. But, what is the reason for this? Can we really obtain the same combinations if we dispense with that column? Let's see that it is easy to answer affirmatively to both questions.

Using the dependence relation (4.3) it is easy to verify that any combination of the four columns can also be obtained using only the first, second and fourth columns. For example, $\begin{pmatrix} 13\\-1\\-11\\1 \end{pmatrix}$ is a combination of the four columns but using (4.3) we have that

$$\begin{pmatrix} 13\\-1\\-11\\1 \end{pmatrix} = 2 \begin{pmatrix} 2\\0\\-1\\1 \end{pmatrix} - 1 \begin{pmatrix} 1\\2\\1\\1 \end{pmatrix} + 2 \underbrace{\begin{pmatrix} 5\\-1\\-7\\-3 \end{pmatrix}}_{\substack{\text{We substitute}\\ \text{using (4.3)}}} +3 \begin{pmatrix} 0\\1\\2\\2 \end{pmatrix} \Rightarrow$$

$$\begin{pmatrix} 13\\-1\\-11\\1 \end{pmatrix} = 2 \begin{pmatrix} 2\\0\\-1\\1 \end{pmatrix} - 1 \begin{pmatrix} 1\\2\\1\\1 \end{pmatrix} + 2 \overbrace{\left(2 \begin{pmatrix} 2\\0\\-1\\1 \end{pmatrix} + 1 \begin{pmatrix} 1\\2\\1\\1 \end{pmatrix} - 3 \begin{pmatrix} 0\\1\\2\\2 \end{pmatrix} \right)}^{\downarrow} +3 \begin{pmatrix} 0\\1\\2\\2 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \begin{pmatrix} 13\\-1\\-11\\1 \end{pmatrix} = 6 \begin{pmatrix} 2\\0\\-1\\1 \end{pmatrix} + \begin{pmatrix} 1\\2\\1\\1 \end{pmatrix} - 3 \begin{pmatrix} 0\\1\\2\\2 \end{pmatrix}.$$

Therefore, it is sufficient to use the first, second and fourth columns. The fact that we can express the third column as a linear combination of the others has allowed us to eliminate it from the linear combination. Actually, this same argument is valid for any combination of the four columns and consequently

$$\underbrace{\left\langle \begin{pmatrix}2\\0\\-1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\1\\1\end{pmatrix}, \begin{pmatrix}5\\-1\\-7\\-3\end{pmatrix}, \begin{pmatrix}0\\1\\2\\2\end{pmatrix} \right\rangle}_{\text{All combinations obtained using the four columns,}} = \underbrace{\left\langle \begin{pmatrix}2\\0\\-1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\2\\2\end{pmatrix} \right\rangle}_{\text{can also be obtained if we remove the third}}.$$

In short, the third column is superfluous when obtaining linear combinations and we could eliminate it,

$$\left\langle \begin{pmatrix}2\\0\\-1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\1\\1\end{pmatrix}, \cancel{\begin{pmatrix}5\\-1\\-7\\-3\end{pmatrix}}, \begin{pmatrix}0\\1\\2\\2\end{pmatrix} \right\rangle.$$

Let's see next some complementary notes about dependence, independence and linear combinations.

Remark.

★ Note that the notation $\langle v_1, v_2, \dots, v_m\rangle$ that we saw in **Definition 124** is used on occasions to abbreviate writing. Thus, instead of writing

'Let v be an element of $\mathbb{R}^3$ that is a linear combination of $(1, 2, -1)$ and $(2, 2, 1)$'

we will put

'Let $v \in \langle(1, 2, -1), (2, 2, 1)\rangle$'.

★ The zero tuple can always be obtained as a linear combination of any tuples $v_1, v_2, \dots, v_m$ (that is, it always holds that $0 \in \langle v_1, v_2, \dots, v_m\rangle$). To do this it is enough to take all the coefficients of the linear combination as zero,

$$0 = 0v_1 + v_2 + \cdots + 0v_m.$$

This way of obtaining the zero tuple, being the simplest, is called 'trivial'.

Example 127. Given the rows $\begin{pmatrix}2 & 3 & 4\end{pmatrix}$ and $\begin{pmatrix}4 & 3 & 9\end{pmatrix}$, we can obtain the zero row (in this case the zero row will be $0_{1\times 3} = \begin{pmatrix}0 & 0 & 0\end{pmatrix}$) by taking equal to zero the two coefficients of the linear combination in the form

$$\begin{pmatrix}0 & 0 & 0\end{pmatrix} = 0\begin{pmatrix}2 & 3 & 4\end{pmatrix} + 0\begin{pmatrix}4 & 3 & 9\end{pmatrix}.$$

This would be the trivial way to obtain the zero tuple.

★ As a consequence of the previous comment, any set of tuples that contains the tuple 0 will always be linearly dependent. Indeed, we know that the tuple 0 can always be obtained as a linear combination of the others and therefore the set of tuples must be dependent.

Example 128. Without needing to perform any calculation we know that the tuples $(1, 2, -1), (2, 1, 1), (0, 0, 0)$ are dependent since one of them is the zero tuple that can be obtained as a linear combination of the others in the form

$$(0, 0, 0) = 0(1, 2, -1) + 0(2, 1, 1).$$

★ Given several tuples $v_1, v_2, \dots, v_m$, in general, the set of their linear combinations, $\langle v_1, v_2, \dots, v_m\rangle$ always has infinitely many elements. That is, there is an infinite amount of linear combinations that can be obtained from certain tuples.

Example 129. If we take the tuples $(2,1)$ and $(3,2)$ we can obtain various linear combinations from them and all will be of the form

$$a_1(2,1) + a_2(3,2),$$

where $a_1, a_2 \in \mathbb{R}$ are the coefficients of the combination. By giving different values to a_1 and a_2 we will obtain their different combinations. Thus for example:

- Taking $a_1 = 3$ and $a_2 = -1$: $3(2,1) + (-1)(3,2) = (3,1)$.
- Taking $a_1 = 2$ and $a_2 = 2$: $2(2,1) + 2(3,2) = (10,6)$.
- Taking $a_1 = 1$ and $a_2 = 0$: $1(2,1) + 0(3,2) = (2,1)$.
- Taking $a_1 = 0$ and $a_2 = 1$: $0(2,1) + 1(3,2) = (3,2)$.
- Taking $a_1 = 0$ and $a_2 = 0$: $0(2,1) + 0(3,2) = (0,0)$.

We have that all the combinations thus obtained are elements of $\langle (2,1), (3,2) \rangle$. That is,

$$(3,1), (10,6), (2,1), (3,2), (0,0) \in \langle (2,1), (3,2) \rangle.$$

In general we have that

$$\langle (2,1), (3,2) \rangle = \underbrace{\{\underbrace{a_1(2,1) + a_2(3,2)}_{\substack{\text{combinations of} \\ (2,1),(3,2)}} : \underbrace{a_1, a_2 \in \mathbb{R}}_{\substack{\text{The coefficients take} \\ \text{different real} \\ \text{values}}} \}}_{\substack{\text{Gathering all these combinations we obtain} \\ \text{all of } \langle (2,1),(3,2) \rangle}}.$$

As in the previous example, in general we have that

$$\langle v_1, v_2, \ldots, v_m \rangle = \underbrace{\{\underbrace{a_1 v_1 + a_2 v_2, \cdots + a_m v_m}_{\substack{\text{Combinations of} \\ v_1,\ldots,v_m}} : \underbrace{a_1, a_2, \ldots, a_m \in \mathbb{R}}_{\substack{\text{The coefficients take} \\ \text{different real} \\ \text{values}}}\}}_{\substack{\text{Gathering all these combinations we obtain} \\ \text{all of } \langle v_1,\ldots,v_m \rangle}}.$$

There is a unique case in which the set of linear combinations is not infinite and has a single element. This case is when the only tuple we have to obtain linear combinations is the zero tuple. Then, the only tuple we can obtain as a linear combination is the zero tuple itself. That is,

$$\langle 0 \rangle = \{0\}.$$

Example 130. Let us take the tuple 0 of $\mathbb{R}^3$, $(0,0,0)$. We have that

$$\langle (0,0,0) \rangle = \underbrace{\{ \underbrace{a_1(0,0,0)}_{\substack{=(0,0,0) \\ \text{always gives the same}}} : a_1 \in \mathbb{R} \}}_{\substack{\text{No matter how many values we give to } a_1 \\ \text{we only obtain } (0,0,0)}} = \{(0,0,0)\}.$$

In this way we see that $\langle (0,0,0) \rangle$ does not have infinitely many elements as happens in general but only one.

★ It is always true that

$$v_1, v_2, \ldots, v_m \in \langle v_1, v_2, \ldots, v_m \rangle.$$

That is, $v_1, v_2, \ldots, v_m$ can be obtained as linear combinations of the tuples $v_1, v_2, \ldots, v_m$. But this is clear since v_i is achieved by taking zero all the coefficients except the one that multiplies v_i which we take equal to 1:

$$v_i = 0v_1 + 0v_2 + \cdots + 0v_{i-1} + 1v_i + 0v_{i+1} + \cdots + 0v_m.$$

Example 131. We have that

$$(2,-1,2) \in \langle (1,2,1),(2,-1,2),(3,3,1) \rangle$$

since we can obtain $(2,-1,2)$ through the following combination

$$(2,-1,2) = 0(1,2,1) + 1(2,-1,2) + 0(3,3,1)$$

in which the coefficients are 0, 1 and 0. Reasoning equally it is also clear that

$$(1,2,1),(3,3,1) \in \langle (1,2,1),(2,-1,2),(3,3,1) \rangle.$$

★ It is evident that if we have more tuples with them we will also be able to perform more linear combinations. That is, if we have the m tuples $v_1, v_2, \dots, v_m \in \mathbb{R}^n$ and additionally we take q more tuples, $w_1, w_2, \dots, w_q \in \mathbb{R}^n$, it holds that

$$\underbrace{\langle v_1, v_2, \dots, v_m \rangle}_{\substack{\text{All combinations} \\ \text{we can obtain} \\ \text{combining } v_1,\dots,v_m}} \subseteq \underbrace{\langle v_1, v_2, \dots, v_m, w_1, w_2, \dots, w_q \rangle}_{\substack{\text{we can also obtain them} \\ \text{combining } v_1,\dots,v_m,w_1,\dots,w_q}}.$$

This is clear since if v is obtained as a linear combination of $v_1, v_2, \dots v_m$ with the coefficients $a_1, a_2, \dots, a_m$,

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_m v_m,$$

then v is also obtained as a linear combination of $v_1, v_2, \dots, v_m$, $w_1, w_2, \dots, w_q$ taking now as coefficients $a_1, a_2, \dots, a_m$ and $0, 0, \dots 0$,

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_m v_m + 0w_1 + 0w_2 + \cdots + 0w_q.$$

Example 132. Given $(2,3,-1,0),(1,2,1,1) \in \mathbb{R}^4$ and additionally $(1,6,3,1),(0,3,0,1) \in \mathbb{R}^4$, any combination of the first two, for example

$$3(2,3,-1,0) + 2(1,2,1,1) = (8,13,-1,2),$$

can also be written as a combination of the four, for example

$$(8,13,-1,2) = 3(2,3,-1,0) + 2(1,2,1,1) + 0(1,6,3,1) + 0(0,3,0,1).$$

Therefore,

$$\langle \text{(2,3,-1,0),(1,2,1,1)} \rangle \subseteq \langle \text{(2,3,-1,0),(1,2,1,1)}, (1,6,3,1),(0,3,0,1) \rangle.$$

★ If $v_1, v_2, \dots, v_p$ are independent, any subset of tuples we choose from among them are also linearly independent.

Example 133. It is possible to verify that the 4-tuples

$$(1,2,-1,1),(2,1,1,1),(0,-1,1,1),(2,-2,1,1)$$

are independent. In such a case any subset of them that we take will also be independent. For example,

$$(1,2,-1,1),(0,-1,1,1),(2,-2,1,1)$$

are independent tuples.

★ If among the tuples $v_1, v_2, \ldots, v_p$ any of them appears repeated, then said tuples are linearly dependent.

Example 134. The 4-tuples

$$(3, 2, -1, 2),\ (2, 1, 2, 1),\ (3, 2, -1, 2),\ (7, 2, 3, 1)$$

are dependent since one of them appears repeated. To see it it is enough to take one of the tuples that appears repeated, $(3, 2, -1, 2)$, and try to obtain it as a combination of the remaining ones which in this case are $(2, 1, 2, 1)$, $(3, 2, -1, 2)$ and $(7, 2, 3, 1)$. Being repeated, among the remaining ones the tuple we want to obtain will appear again and then it is simple to pose the combination by taking zero all the coefficients except the one corresponding to the repeated tuple:

$$\underbrace{(3, 2, -1, 2)}_{\text{The repeated tuple}} = \underbrace{0(2, 1, 2, 1) + 1\underbrace{(3, 2, -1, 2)}_{\text{appears among the remaining ones}} + 0(7, 2, 3, 1)}_{\text{and therefore we can obtain it easily as a linear combination.}}$$

Aquí está la traducción al inglés del fragmento sobre matrices, cumpliendo estrictamente con todas las indicaciones proporcionadas.

"'latex

4.3.1 Basic techniques for studying linear dependence

We already know what a linear combination is and we are familiar with the concepts of dependence and independence. We now need techniques that allow us to solve problems related to these concepts. Specifically, we need to solve the following problems:

a) Determine whether a given tuple can or cannot be obtained by combining others.

b) Determine whether a set of tuples are dependent or independent.

Let's see how we can do this.

Determining if a tuple is a linear combination of others

We know that a tuple, $w \in \mathbb{R}^n$, is a linear combination of $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$ if we can find numbers $\alpha_1, \alpha_2, \ldots, \alpha_m$ such that

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m.$$

In reality, when $\alpha_1, \alpha_2, \ldots, \alpha_m$ are unknown, the expression above constitutes a system of linear equations whose variables are the coefficients we want to determine. If we can solve this system we can find the numbers we are looking for and w can be obtained as a linear combination; otherwise, it will not be possible.

Let's see this better in the following examples.

Examples 135.

1) Determine if the tuple

$$N_1 = \begin{pmatrix} 1 \\ 3 \\ 5 \\ 0 \\ -1 \\ 10 \\ 4 \end{pmatrix}$$

is a linear combination of the tuples N_C and N_M from **Example 123**. N_1 will be a linear combination if we can find the numbers $\alpha, \beta \in \mathbb{R}$ such that

$$N_1 = \alpha N_C + \beta N_M.$$

If we substitute the value of each tuple and perform the matrix operations appearing in the expression, we will have

$$\begin{pmatrix} 1 \\ 3 \\ 5 \\ 0 \\ -1 \\ 10 \\ 4 \end{pmatrix} = \alpha \underbrace{\begin{pmatrix} 7 \\ 8 \\ 10 \\ 6 \\ 4 \\ 20 \\ 9 \end{pmatrix}}_{=N_C} + \beta \underbrace{\begin{pmatrix} 6 \\ 5 \\ 5 \\ 6 \\ 5 \\ 10 \\ 5 \end{pmatrix}}_{=N_M} \quad \begin{matrix} \Rightarrow \\ \text{Performing} \\ \text{the operations} \end{matrix} \quad \begin{pmatrix} 1 \\ 3 \\ 5 \\ 0 \\ -1 \\ 10 \\ 4 \end{pmatrix} = \begin{pmatrix} 7\alpha + 6\beta \\ 8\alpha + 5\beta \\ 10\alpha + 5\beta \\ 6\alpha + 6\beta \\ 4\alpha + 5\beta \\ 20\alpha + 10\beta \\ 9\alpha + 5\beta \end{pmatrix}$$

and if we now equate row by row we arrive at

$$\begin{cases} 7\alpha + 6\beta = 1 \\ 8\alpha + 5\beta = 3 \\ 10\alpha + 5\beta = 5 \\ 6\alpha + 6\beta = 0 \\ 4\alpha + 5\beta = -1 \\ 20\alpha + 10\beta = 10 \\ 9\alpha + 5\beta = 4 \end{cases}.$$

As we said, when the values of the coefficients are unknown, setting up the linear combination yields a system of linear equations. We must now try to solve this system to answer the question posed. In this case, it is easy to find the solution. For example, if we subtract the second equation from the third we get that $\alpha = 1$ and then substituting into any other equation we can solve for $\beta = -1$. We can check that all equations are satisfied for these values of α and β. Therefore, we have calculated the coefficients α and β that we need and we now know that N_1 can be written as

$$N_1 = N_C - N_M.$$

Thus $N \in \langle N_C, N_M \rangle$, that is, N can be obtained by combining N_C and N_M as indicated in this last relation of dependence.

Note that the problem reduces to finding the solution of a system of linear equations.

Let us now study the same problem for the tuple

$$N_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Repeating the same steps, we must again find the coefficients α and β such that

$$N_2 = \alpha N_C + \beta N_M \Rightarrow \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \alpha \underbrace{\begin{pmatrix} 7 \\ 8 \\ 10 \\ 6 \\ 4 \\ 20 \\ 9 \end{pmatrix}}_{=N_C} + \beta \underbrace{\begin{pmatrix} 6 \\ 5 \\ 5 \\ 6 \\ 5 \\ 10 \\ 5 \end{pmatrix}}_{=N_M} \Rightarrow \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7\alpha + 6\beta \\ 8\alpha + 5\beta \\ 10\alpha + 5\beta \\ 6\alpha + 6\beta \\ 4\alpha + 5\beta \\ 20\alpha + 10\beta \\ 9\alpha + 5\beta \end{pmatrix} \Rightarrow \begin{cases} 7\alpha + 6\beta = 1 \\ 8\alpha + 5\beta = 1 \\ 10\alpha + 5\beta = 1 \\ 6\alpha + 6\beta = 1 \\ 4\alpha + 5\beta = 1 \\ 20\alpha + 10\beta = 1 \\ 9\alpha + 5\beta = 1 \end{cases}.$$

We must now determine if this system has a solution or not. If we subtract the second equation from the third we get that $\alpha = 0$. If we now solve for β in the first equation we get $\beta = 1/6$. On the other hand, if we solve for it in the second equation, $\beta = 1/8$. Since β cannot take two different values simultaneously, we conclude that the system has no solution and, consequently, it is not possible to find the coefficients α and β that we need. Therefore, we finally have that

$$N_2 \notin \langle N_C, N_M \rangle$$

and the tuple N_2 cannot be obtained by combining N_C and N_M.

2) Check if $(3,3,1) \in \langle (1,2,1),(1,1,1),(2,1,1) \rangle$.

Again we must find the coefficients, in this case three, needed to form the combination that produces the tuple $(3,3,1)$,

$$(3,3,1) = \alpha(1,2,1) + \beta(1,1,1) + \gamma(2,1,1).$$

Performing the operations and equating we have

$$(3,3,1) = (\alpha + \beta + 2\gamma, 2\alpha + \beta + \gamma, \alpha + \beta + \gamma) \Rightarrow \begin{cases} \alpha + \beta + 2\gamma = 3 \\ 2\alpha + \beta + \gamma = 3 \\ \alpha + \beta + \gamma = 1 \end{cases}.$$

To solve the system,

$$\begin{cases} \text{subtracting the last equation from the first} & \Rightarrow \quad \gamma = 2, \\ \text{subtracting the last equation from the second} & \Rightarrow \quad \alpha = 2, \\ \text{substituting and solving in the third} & \Rightarrow \quad \beta = -3. \end{cases}$$

Therefore

$$(3,3,1) = 2(1,2,1) - 3(1,1,1) + 2(2,1,1)$$

and $(3,3,1) \in \langle (1,2,1),(1,1,1),(2,1,1) \rangle$.

3) Study whether $(1,0,0) \in \langle (1,2,1),(1,1,1),(2,3,2) \rangle$.

Repeating the process we will again obtain a system of linear equations,

$$(1,0,0) = \alpha(1,2,1) + \beta(1,1,1) + \gamma(2,3,2) \Rightarrow \begin{cases} \alpha + \beta + 2\gamma = 1 \\ 2\alpha + \beta + 3\gamma = 0 \\ \alpha + \beta + 2\gamma = 0 \end{cases}$$

and it is evident that the first and the last equation cannot be satisfied at the same time since $\alpha+\beta+2\gamma$ cannot simultaneously be 1 and 0. Consequently, this system has no solution and we cannot find the coefficients α, β and γ. Therefore,

$$(1,0,0) \notin \langle (1,2,1),(1,1,1),(2,3,2) \rangle.$$

Study of linear dependence and independence

Given several tuples $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$ we know that it is always possible to combine them to trivially obtain the tuple 0 of $\mathbb{R}^n$ (which is the n-tuple with all its entries equal to 0, $0 = (0,0,\ldots,0) \in \mathbb{R}^n$). To do this, it is enough to form the linear combination of them with all coefficients equal to zero since clearly

$$0v_1 + 0v_2 + \cdots + 0v_m = 0.$$

But, is this the only way to obtain the tuple 0 as a combination of $v_1, v_2, \ldots, v_m$ or will there be others? In fact, this question is the key to determining whether the tuples $v_1, v_2, \ldots, v_m$ are independent or not.

To justify this last point, let's return to **Example 123** where we were studying the motor vehicle fleet in different cities and take the tuples N_C, N_M and N_R which respectively contained the information corresponding to the number of cars, motorcycles, and wheels circulating in each city. Since

$$N_R = 4N_C + 2N_M,$$

we know that these three tuples, $\{N_C, N_M, N_R\}$, are dependent since one of them can be obtained by combining the others. If we move all the terms in this last equality to the same side we have

$$4N_C + 2N_M - N_R = 0,$$

where $0 = (0, 0, 0, 0, 0, 0, 0)$ is the 7-tuple zero. In this way, since they are dependent, we can find a way to obtain the tuple 0 different from the trivial one.

This idea, which we have just illustrated with this example, is precisely captured in the following property that outlines the basic technique for studying the dependence or independence of a set of tuples.

Property 136. *Consider the tuples $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$. Then if the only values of the numbers $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathbb{R}$ for which we obtain*

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m = 0$$

are $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$, said tuples are independent. Otherwise, the tuples will be dependent.

Let's see this more clearly in the following examples.

Examples 137.

1) Let us study whether the tuples $(2, 3, 1), (4, 6, 4), (4, 6, 3) \in \mathbb{R}^3$ are dependent or independent. To do this we must determine for which numbers α_1, α_2 and α_3 we have

$$\alpha_1(2, 3, 1) + \alpha_2(4, 6, 4) + \alpha_3(4, 6, 3) = (0, 0, 0).$$

First we will perform the matrix operations indicated in this equality, which leads us to

$$(2\alpha_1 + 4\alpha_2 + 4\alpha_3, 3\alpha_1 + 6\alpha_2 + 6\alpha_3, \alpha_1 + 4\alpha_2 + 3\alpha_3) = (0, 0, 0)$$

and now equating both members we finally obtain the system of linear equations with three variables and three equations,

$$\left\{\begin{array}{l} 2\alpha_1 + 4\alpha_2 + 4\alpha_3 = 0 \\ 3\alpha_1 + 6\alpha_2 + 6\alpha_3 = 0 \\ \alpha_1 + 4\alpha_2 + 3\alpha_3 = 0 \end{array}\right. .$$

Solving this system we must determine the possible values for α_1, α_2 and α_3. In this case,

$$\left\{\begin{array}{lcl} 2\alpha_1 + 4\alpha_2 + 4\alpha_3 = 0 & \xrightarrow[\text{dividing by 2}]{} & \cancel{\alpha_1 + 2\alpha_2 + 2\alpha_3 = 0} \\ 3\alpha_1 + 6\alpha_2 + 6\alpha_3 = 0 & \xrightarrow[\text{dividing by 3}]{} & \alpha_1 + 2\alpha_2 + 2\alpha_3 = 0 \\ \alpha_1 + 4\alpha_2 + 3\alpha_3 = 0 & & \alpha_1 + 4\alpha_2 + 3\alpha_3 = 0 \end{array}\right. ,$$

where we have eliminated the first equation because it coincides with the second. Now the system becomes

$$\left\{\begin{array}{lcl} \alpha_1 + 2\alpha_2 + 2\alpha_3 = 0 & \xrightarrow[\text{equation 2 minus equation 1}]{} & \alpha_3 = -2\alpha_2 \\ \alpha_1 + 4\alpha_2 + 3\alpha_3 = 0 & \xrightarrow[\text{substituting}]{} & \alpha_1 = 2\alpha_2 \end{array}\right. .$$

This is therefore a compatible indeterminate system that will have infinitely many solutions. For each value of α_2 we can calculate α_1 and α_3 using the equations we have obtained. Since we have infinitely many solutions we can find expressions for the zero tuple different from the trivial one and directly **Property 136** tells us that these tuples are dependent.

Although we already know that the tuples are dependent, as an example, we can obtain a specific solution of the system and its corresponding linear combination. If we take $\alpha_2 = 1$ we obtain, using the previous equations, the solution

$$\alpha_1 = 2, \ \alpha_2 = 1, \ \alpha_3 = -2,$$

which leads us to the following non-trivial expression of the zero tuple,

$$2(2,3,1) + (4,6,4) - 2(4,6,3) = (0,0,0).$$

In fact, from this non-trivial combination we can deduce that the tuples are dependent simply using the definition of dependence since it allows us to express one of the tuples as a linear combination of the others. For example,

$$(0,0,0) = \boxed{2}\,(2,3,1) + 1\,(4,6,4) + (-2)\,(4,6,3) \Rightarrow (0,0,0) - \boxed{2}\,(2,3,1) = 1\,(4,6,4) + (-2)\,(4,6,3)$$

$$\Rightarrow (2,3,1) = \frac{1}{-\boxed{2}}(4,6,4) + \frac{2}{-\boxed{2}}(4,6,3)$$

We see in this way that $(2,3,1)$ is a linear combination of $(4,6,4)$ and $(4,6,3)$ and we have again the dependence.

2) Let us check if the tuples $(1,1,3), (2,0,-1), (-1,0,1)$ are dependent or independent.

To do this, we will use **Property 136**. Suppose that a certain linear combination of the tuples yields the zero tuple,

$$a_1(1,1,3) + a_2(2,0,-1) + a_3(-1,0,1) = (0,0,0).$$

We must determine the coefficients of this combination, a_1, a_2 and a_3. To do this we will solve the system of linear equations that arises when we perform the operations indicated in the linear combination and equate. Let's see it,

$$a_1(1,1,3) + a_2(2,0,-1) + a_3(-1,0,1) = (0,0,0)$$

$$\Downarrow$$

$$(a_1, a_1, 3a_1) + (2a_2, 0, -a_2) + (-a_3, 0, a_3) = (0,0,0)$$

$$\Downarrow$$

$$(a_1 + 2a_2 - a_3, a_1, 3a_1 - a_2 + a_3) = (0,0,0)$$

$\Downarrow$ For two tuples to be equal, each of their components must be equal.

$$\begin{cases} a_1 + 2a_2 - a_3 = 0 \\ a_1 = 0 \\ 3a_1 - a_2 + a_3 = 0 \end{cases}$$

The coefficients a_1, a_2 and a_3 must satisfy the equations of the previous system. Now, if we solve the system we easily obtain that

$$a_1 = 0, \quad a_2 = 0, \quad a_3 = 0$$

and therefore the coefficients of the combination are all necessarily equal to zero. In this way, we verify that in this case the only way to obtain the zero tuple is to take all the coefficients of the combination as zero and we deduce that the tuples are independent.

3) Study the dependence and independence of the tuples $(2,0,0,1)$, $(0,1,0,-1)$, $(4,-1,0,3)$, $(0,1,1,0)$.

We will use the same technique as in the previous section, equating to zero a linear combination of the tuples and obtaining the corresponding system of equations:

$$x(2,0,0,1)+y(0,1,0,-1)+z(4,-1,0,3)+w(0,1,1,0)=(0,0,0,0)$$
$$\Downarrow$$
$$(2x,0,0,x)+(0,y,0,-y)+(4z,-z,0,3z)+(0,w,w,0)=(0,0,0,0)$$
$$\Downarrow$$
$$(2x+4z,y-z+w,w,x-y+3z)=(0,0,0,0)$$
$$\Downarrow$$
$$\left\{\begin{array}{l} 2x+4z=0 \\ y-z+w=0 \\ w=0 \\ x-y+3z=0 \end{array}\right. .$$

If we solve the system we observe that we can only solve for three of the variables in the form

$$\left\{\begin{array}{l} x=-2z \\ y=z \\ w=0 \end{array}\right. .$$

It is easy to realize that this is a system with infinitely many solutions that we can obtain by giving different values to z in the above equalities. Since we have solutions different from the trivial one, directly, via **Property 136**, we deduce that the tuples are dependent.

4) Consider the n elements $e_1, e_2, e_3, \ldots, e_n$ of $\mathbb{R}^n$ defined as follows:

$$\begin{array}{rcl} e_1 & = & (1,0,0,0,\ldots,0,0), \\ e_2 & = & (0,1,0,0,\ldots,0,0), \\ e_3 & = & (0,0,1,0,\ldots,0,0), \\ & \vdots & \\ e_n & = & (0,0,0,0,\ldots,0,1). \end{array}$$

It is easy to see that they are independent using again **Property 136**. To do this, suppose we obtain the tuple 0 of $\mathbb{R}^n$ (i.e., $0=(0,0,0,\ldots,0)$) as a linear combination of e_1, e_2, $\ldots$, e_n in the form

$$\alpha_1 e_1+\alpha_2 e_2+\cdots+\alpha_n e_n=0$$

and let's see that in such a case the only possibility is that all the coefficients α_1, α_2, $\ldots$, α_n are equal to zero. We have that

$$\begin{gathered} \alpha_1 e_1+\alpha_2 e_2+\cdots+\alpha_n e_n=0 \\ \Rightarrow \alpha_1(1,0,0,\ldots,0)+\alpha_2(0,1,0,\ldots,0)+\ldots\alpha_n(0,0,0,\ldots,1)=(0,0,0,\ldots,0) \\ \Rightarrow(\alpha_1,0,0,\ldots,0)+(0,\alpha_2,0,\ldots,0)+(0,0,0,\ldots,\alpha_n)=(0,0,0,\ldots,0) \\ \Rightarrow(\alpha_1,\alpha_2,\ldots,\alpha_n)=(0,0,\ldots,0) \\ \Rightarrow\left\{\begin{array}{l} \alpha_1=0, \\ \alpha_2=0, \\ \quad\vdots \\ \alpha_n=0. \end{array}\right. \end{gathered}$$

Therefore all the coefficients α_1, α_2, $\ldots$, α_n are necessarily zero and as a consequence $e_1, e_2, \ldots, e_n$ are independent.

The tuples e_1, e_2,$\ldots$, e_n of $\mathbb{R}_n$ are called coordinate n-tuples of $\mathbb{R}^n$ and we will see that they play an important role within matrix theory. For each set $\mathbb{R}^2$, $\mathbb{R}^3$, etc. we have a different set of coordinate n-tuples. Let's look at some examples:

- In $\mathbb{R}^2$ the coordinate 2-tuples are 2:

$$e_1 = (1, 0) \quad \text{and} \quad e_2 = (0, 1).$$

We also know that e_1 and e_2 are independent of each other.

- In $\mathbb{R}^3$ the coordinate 3-tuples will be 3:

$$e_1 = (1, 0, 0), \; e_2 = (0, 1, 0) \quad \text{and} \quad e_3 = (0, 0, 1).$$

These three tuples are independent of each other.

- In $\mathbb{R}^4$ the 4 coordinate tuples are

$$e_1 = (1, 0, 0, 0), \; e_2 = (0, 1, 0, 0), \; e_3 = (0, 0, 1, 0), \quad \text{and} \quad e_4 = (0, 0, 0, 1).$$

Again, these four tuples form a linearly independent set.

In the same way we have the coordinate n-tuples for any set $\mathbb{R}^n$.

Note that the notation e_1, e_2, etc. is ambiguous. Depending on whether we are in $\mathbb{R}^2$, $\mathbb{R}^3$ or $\mathbb{R}^4$, e_1 or e_2 will mean different things. It is similar to what happens with the tuple 0 which will be $0 = (0, 0)$ for $\mathbb{R}^2$ or $0 = (0, 0, 0)$ for $\mathbb{R}^3$. It is necessary to be careful with the problems posed by this ambiguity and as always, determine based on the context of each problem the appropriate option.

4.3.2 Rank of a Matrix

In this section we will introduce the concept of the rank of a matrix and we will describe techniques to compute it easily. Through the properties we will see for the rank we will be able to study the dependence and independence of sets of tuples quickly.

Given any matrix we can consider its row or column tuples. The analysis of these row or column tuples allows us to define the concept of the rank of a matrix.

Definition 138. Given the matrix $A = (v_1|v_2|\dots|v_n) \in \mathcal{M}_{m\times n}$ whose column tuples are v_1, v_2, …, v_n, we call the rank of the matrix A, and denote it by

$$\text{rango}(A) \quad \text{or} \quad \text{r}(A),$$

the size of the largest subset of independent tuples that we can find among the column tuples v_1, v_2, …, v_n. By definition, we will say that the rank of the matrix $0_{m\times n}$ is 0.

Examples 139.
1) Let's take

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Its column tuples are

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \; \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The first two correspond to the coordinate tuples e_1 and e_2 of $\mathbb{R}^3$ and the last two are equal to the tuple 0. To calculate the rank we have to obtain independent sets from the column tuples and keep the largest of these sets (the one with the largest number of elements). The 3-tuples coordinates of $\mathbb{R}^3$, which are e_1, e_2

and e_3, form an independent set and any subset of it remains independent so $\{e_1, e_2\}$ is linearly independent. Therefore, taking the first two columns of A, we have the set

$$\{e_1, e_2\}$$

which is independent and has size 2. We can ask ourselves if there will exist some set of independent columns with more elements. However, we know that any set that contains the tuple 0 will always be dependent. Therefore the last two columns can never be part of any independent set. From this we deduce that any set of column tuples larger than $\{e_1, e_2\}$ should include at least one of the tuples 0 and would not be independent. In this way, the largest set of column tuples independent that we can achieve is $\{e_1, e_2\}$ which has size 2 and consequently

$$\text{rango}(A) = 2.$$

2) Let's take the matrix

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Its column tuples,

$$(1, 0, 0, 0), \ (0, 1, 0, 0), \ (0, 0, 1, 0)$$

coincide with the first three coordinate tuples of $\mathbb{R}^4$, e_1, e_2 and e_3. But any subset of the coordinate tuples forms an independent subset and therefore the column tuples of the matrix B are independent. The largest linearly independent set that we can obtain from these three columns will therefore consist of taking them all and for that reason it will have size 3. Consequently,

$$\text{rango}(B) = 3.$$

3) Using the same arguments as in the two previous examples we can calculate at a glance the rank of some simple matrices:

- $\text{rango}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = 3.$
- $\text{rango}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 4.$
- $\text{rango}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1.$
- $\text{rango}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = 2.$

4) In **Definition 138** it is established that the rank of the zero matrix is always 0. Some examples of this are:

- $\text{rango}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 0.$

- rango $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$.

In reality, the column tuples of the zero matrix are all of them equal to the corresponding zero tuple and therefore we can never form any independent set with them. Hence the rank of these matrices is established as 0.

In the previous examples we see that it is possible to calculate the rank of simple matrices in which all elements are zero except for some ones arranged diagonally. Schematically we represent all of them in the form

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right),$$

which designates a matrix with all its elements zero except for r ones arranged diagonally. Taking into account the examples seen, the following property is evident:

Property 140. *Given $r, n, m \in \mathbb{N}$:*

i) rango $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) = r$.

ii) $\text{rango}(I_r) = r$.

iii) $\text{rango}(0_{n\times m}) = 0$.

In other words, to calculate the rank of a matrix with all its elements zero except for ones on the diagonal, we only need to count the number of ones that appear in it. For example, the identity matrix of order r will always have rank r since it has r ones arranged diagonally.

The question is that so far we only know how to calculate the rank of matrices of the type $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$ and obviously not all matrices are of this form. If we have any matrix we can try to transform it into one of the type $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$ and then calculate its rank. Of course, the transformation we apply must respect the value of the rank. Let's see the following example:

Example 141. The matrix $A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ is not of the type $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$, however it is sufficient to modify the order of the columns to obtain:

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{Reordering columns}]{} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now, the rank is the size of the largest set of independent columns and evidently the initial matrix, A, and the one obtained after applying the transformation have the same columns. Therefore, it is clear that the rank of both must coincide. But the transformed matrix is clearly of the type $A = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$ and its rank is evidently 3 (three ones on the diagonal). As a consequence

$$\text{r}(A) = 3.$$

Changing the order of the columns of a matrix does not alter the rank and in the previous example this property has allowed us to calculate the rank of the matrix A which we initially did not know. Unfortunately if we start from an arbitrary matrix, we will not always be able to arrive at one of the type $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$ by modifying the order of the columns. However, we can ask ourselves the following two questions:

a) Are there more transformations like this one that modify the matrix but not the value of the rank?

b) Assuming the answer to question a) is affirmative: Will the available transformations allow us to convert any matrix into one of the type $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$?

In the remainder of this section we will see that in both cases we have an affirmative answer. We begin by seeing in the following property three types of transformations that do not modify the rank. With this we solve question a).

Property 142. *Given the matrix $A \in \mathcal{M}_{m\times n}$, it holds that:*

i) If we modify the order of the rows or columns of A, the resulting matrix has the same rank as A.

ii) If we multiply one of the rows or columns of A by a number different from zero, the resulting matrix has the same rank as A.

iii) If we add to one column (respectively row) another column (respectively row) multiplied by a number, the resulting matrix has the same rank as A.

In the previous property, we affirm that there are three types of transformations that we can apply to a matrix without the rank varying. These transformations are what are usually called elementary operations on the matrix. Specifically we have:

Definition 143. Given a matrix $A = (a_{ij})_{m\times n} \in \mathcal{M}_{m\times n}$ we call an elementary operation on A any of the following transformations:

- Multiply a row or column by a non-zero number.
- Modify the order of the rows or columns.
- Add to one column (respectively row) another column (resp. row) multiplied by any number.

Then **Property 142** tells us that elementary operations are transformations that preserve the rank of a matrix.

In what follows we will use the following nomenclature to describe the elementary operations we perform on a matrix:

a) When we multiply the i-th column by a number k we indicate it by "kCi".

b) When we interchange column i with column j we indicate it by "Ci↔Cj".

c) When we add to column i the column j multiplied by a number k we denote it by "Ci=Ci+kCj".

d) The same operations for rows are denoted using the letter "F" instead of "C".

Examples 144.
1) Let's apply a series of elementary operations to the following matrix:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 0 \end{pmatrix} \xrightarrow[\text{C1}\leftrightarrow\text{C2}]{} \begin{pmatrix} 2 & 1 & 3 \\ -1 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow[\text{F1=F1+2F2}]{} \begin{pmatrix} 1 & 0 & 3 \\ 0 & -1 & 0 \end{pmatrix}$$

$$\xrightarrow[\text{F1}\leftrightarrow\text{F2}]{} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 2 & 3 \end{pmatrix}$$

$$\xrightarrow[\text{3C1}]{} \begin{pmatrix} 3 & 6 & 9 \\ 0 & -1 & 0 \end{pmatrix}$$

The matrices that appear on the right, all of them, have been obtained from $\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 0 \end{pmatrix}$ by applying elementary operations. Since elementary operations preserve the rank, all of them will have the same rank as the initial matrix:

$$\text{rango}\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 0 \end{pmatrix} = \text{rango}\begin{pmatrix} 2 & 1 & 3 \\ -1 & 0 & 0 \end{pmatrix} = \text{rango}\begin{pmatrix} 1 & 0 & 3 \\ 0 & -1 & 0 \end{pmatrix}$$

$$= \text{rango}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 2 & 3 \end{pmatrix} = \text{rango}\begin{pmatrix} 3 & 6 & 9 \\ 0 & -1 & 0 \end{pmatrix}.$$

2) To calculate the rank of the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ -4 & 1 & 0 \end{pmatrix}$$

we can apply elementary operations to it to try to transform it into a matrix of the type $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$. Let's see how we can do it:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ -4 & 1 & 0 \end{pmatrix} \xrightarrow[\text{C2}\leftrightarrow\text{C3}]{} \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix} \xrightarrow[\text{F2=F2-2F1}]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[\text{F3=F3+4F1}]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Indeed, the last matrix obtained is of the type $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$ and therefore we know how to calculate its rank. At the same time said matrix has been obtained from A through elementary operations and its rank is the same as that of A. In short, we calculate the rank of A as

$$\text{rango}(A) = \text{rango}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 3.$$

The last example reproduces the method we intend to apply to calculate the rank of a matrix. Given the matrix A we will apply elementary operations to it to try to transform it into a matrix of the type $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$,

$$A \xrightarrow[\text{Elementary operations}]{} \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Then we calculate the rank of A through the matrix obtained as

$$\text{rango}(A) = \text{rango}\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) = r.$$

However, in this whole process the question **b)** that we formulated on page 155 remains unanswered. That is: Will we always be able to find the adequate elementary operations that transform A into $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$? The affirmative answer is given by the following property. The proof of the property is based on the application of the Gaussian elimination method that we describe below.

The Gaussian elimination method

The Gaussian elimination method or Gaussian elimination allows us to reduce, through elementary operations, any matrix to a matrix with ones on the diagonal and the rest of the elements zero. It is what is called an iterative method. That is, it is based on the repeated application of the same steps. These steps are the six we will see below and they will be the ones we will always apply to calculate the rank of a matrix.

We will see the steps to follow while reproducing them on a concrete example. So, these are the steps of the Gaussian elimination method:

1) We select in the matrix a row or column in which there exist at least two non-zero elements. In general we will select the row or column with a larger number of zeros.

2) In the row or column selected in the previous section we choose a non-zero element which we call the 'pivot'. We must take into account the following criteria:

- In the case of manual calculations, the choice of the pivot element with value equal to 1 or -1 can simplify the operations.
- In the case of precise calculations, the best results are obtained by selecting as the pivot the element of the row or column with the largest absolute value.

Example 145. Let's apply the method to the matrix

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & -1 & 2 & 1 \\ \mathbf{2} & \mathbf{0} & \overline{\underline{\mathbf{1}}} & \mathbf{0} \\ 1 & -1 & 2 & 3 \end{pmatrix}.$$

We have selected the third row since it has more zeros than any other row or column. Within the row we take the third element as the pivot, whose value is 1, with the objective of simplifying subsequent calculations.

---

3) We use the pivot to cancel all the elements of the row or column initially selected.

Example 146. Within the selected row there is only one non-zero element different from the pivot. We must perform an operation to eliminate this element

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & -1 & 2 & 1 \\ \mathbf{2} & \mathbf{0} & \overline{\underline{\mathbf{1}}} & \mathbf{0} \\ 1 & -1 & 2 & 3 \end{pmatrix} \xrightarrow[\text{C1=C1-2C3}]{} \begin{pmatrix} 1 & 2 & 0 & 1 \\ -3 & -1 & 2 & 1 \\ \mathbf{0} & \mathbf{0} & \overline{\underline{\mathbf{1}}} & \mathbf{0} \\ -3 & -1 & 2 & 3 \end{pmatrix}.$$

Note that to cancel elements in a row we must perform operations by columns. Similarly, if we had to eliminate elements of a column we would do operations with rows.

---

4) We use the pivot to cancel the elements of the row or column perpendicular to the one we had selected in step 1 that intersects at the height of the pivot.

Example 147. We select the column perpendicular at the height of the pivot:

$$\begin{pmatrix} 1 & 2 & \mathbf{0} & 1 \\ -3 & -1 & \mathbf{2} & 1 \\ \mathbf{0} & \mathbf{0} & \underline{\overline{1}} & \mathbf{0} \\ -3 & -1 & \mathbf{2} & 3 \end{pmatrix}.$$

Since we are going to eliminate elements in a column we will do operations by rows.

$$\begin{pmatrix} 1 & 2 & \mathbf{0} & 1 \\ -3 & -1 & \mathbf{2} & 1 \\ \mathbf{0} & \mathbf{0} & \underline{\overline{1}} & \mathbf{0} \\ -3 & -1 & \mathbf{2} & 3 \end{pmatrix} \xrightarrow[\text{F4=F4-2F3}]{\text{F2=F2-2F3}} \begin{pmatrix} 1 & 2 & \mathbf{0} & 1 \\ -3 & -1 & \mathbf{0} & 1 \\ \mathbf{0} & \mathbf{0} & \underline{\overline{1}} & \mathbf{0} \\ -3 & -1 & \mathbf{0} & 3 \end{pmatrix}$$

---

5) Whenever there remains some row or column with more than one non-zero element we return to step 1.

Example 148. It is evident that in the resulting matrix from the previous example there remain rows and columns with more than one non-zero element. We then return to step 1 selecting one of these rows or columns. We will take the second column and as pivot its second element and apply steps 3 and 4:

$$\begin{pmatrix} 1 & \mathbf{2} & 0 & 1 \\ -3 & \underline{\overline{-1}} & 0 & 1 \\ 0 & \mathbf{0} & 1 & 0 \\ -3 & \mathbf{-1} & 0 & 3 \end{pmatrix} \xrightarrow[\text{F4=F4-F2}]{\text{F1=F1+2F2}} \begin{pmatrix} -5 & \mathbf{0} & 0 & 3 \\ -3 & \underline{\overline{-1}} & 0 & 1 \\ 0 & \mathbf{0} & 1 & 0 \\ 0 & \mathbf{0} & 0 & 2 \end{pmatrix} \xrightarrow[\text{C4=C4+C2}]{\text{C1=C1-3C2}} \begin{pmatrix} -5 & \mathbf{0} & 0 & 3 \\ \mathbf{0} & \underline{\overline{-1}} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{0} & 1 & 0 \\ 0 & \mathbf{0} & 0 & 2 \end{pmatrix}.$$

Since there are still rows and/or columns with more than one non-zero element we return again to step 1. We now take the last row and as pivot the first element:

$$\begin{pmatrix} -5 & 0 & 0 & 3 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \underline{\overline{2}} \end{pmatrix} \rightarrow \begin{pmatrix} -5 & 0 & 0 & \mathbf{3} \\ 0 & -1 & 0 & \mathbf{0} \\ 0 & 0 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \underline{\overline{2}} \end{pmatrix} \xrightarrow{\text{F1=F1-}\frac{3}{2}\text{F4}} \begin{pmatrix} -5 & 0 & 0 & \mathbf{0} \\ 0 & -1 & 0 & \mathbf{0} \\ 0 & 0 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \underline{\overline{2}} \end{pmatrix}.$$

In the resulting matrix there are no rows or columns with more than one non-zero element. We will then proceed to the last step.

---

6) If necessary, the rows or columns are reordered to bring the matrix to diagonal form. If on the main diagonal appear non-zero elements different from 1 we can appropriately divide the corresponding row or column to transform them into 1.

Example 149. The matrix, as it was left in the last example, is already in diagonal form so it is not necessary to reorder rows or columns. Since on the main diagonal there are non-zero elements different from 1 we will appropriately divide the corresponding rows:

$$\begin{pmatrix} -5 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \xrightarrow[\text{F2=}\frac{1}{-1}\text{F2}]{\text{F1=}\frac{1}{-5}\text{F1}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = I_4.$$

Finally we have obtained a matrix of the type $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$. We can now calculate the rank of the initial matrix A:

$$\text{rango}\begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & -1 & 2 & 1 \\ 2 & 0 & 1 & 0 \\ 1 & -1 & 2 & 3 \end{pmatrix} = \text{rango}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 4.$$

Property 142 and the Gaussian elimination method answer questions **a)** and **b)** that we formulated on page 155. These two properties, especially the Gaussian elimination method, together with **Property 140** provide us with a mechanism to calculate the rank. In this way, the idea for calculating the rank of a matrix A is based on the following steps:

1. We apply the Gaussian elimination method to find a list of elementary operations, let's call it L, that transforms A as follows:
$$A \xrightarrow[\text{Operations } L]{} \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$
2. We use **Property 142** which guarantees us that
$$\text{rango}(A) = \text{rango}\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$
3. Finally, we apply **Property 140** to perform the calculation:
$$\text{rango}(A) = \text{rango}\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) = r.$$

Row rank and other basic properties

We will now consider a question that remained implicitly pending when in **Definition 138** we established that the rank of a matrix A is the size of the largest independent set of columns of it. The question is whether we could have defined the rank by taking the largest possible set of independent rows instead of working with columns.

Closely related to the above is this other question. If we take the rows of A and put them in column form we obtain the transpose matrix A^t. To calculate the rank of A^t we will take into account its columns which are actually the rows of A. We can then ask what relationship exists between rango(A) and rango(A^t).

The following property uses the results of this section to solve these problems.

Theorem 150. *Given a matrix $A \in \mathcal{M}_{m\times n}$ it holds that:*

i) rango(A) = rango(A^t).

ii) The rank of A is the size of the largest independent set formed by row tuples of the matrix A.

Proof. We begin proving *i)*. If rango$(A) = r$ we will be able to find a set of elementary operations, which we will call L, such that

$$A \xrightarrow[\text{Operations } L]{} \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

If in the set of elementary operations L all the operations by rows are changed to operations by columns and vice versa, we will obtain another different set of operations that we can call L^t. Then it is evident that

$$A^t \xrightarrow[\text{Operations } L^t]{} \left(\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)_{m\times n}\right)^t = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)_{n\times m}.$$

Consequently,

$$\text{rango}(A^t) = \text{rango}\left(\begin{array}{c|c} I_r & 0 \\ 0 & 0 \end{array}\right)_{n\times m} = r = \text{rango}(A).$$

Once we have proven *i)*, proving *ii)* is easy since calculating the size of the largest independent set of rows of A is the same as calculating the size of the largest set of independent columns of A^t which by definition is $\text{rango}(A^t)$. But $\text{rango}(A^t) = \text{rango}(A)$ and therefore the size of the largest independent set of rows of A will be $\text{rango}(A)$. □

An immediate consequence of the last theorem is the fact that the rank of a matrix will be less than or equal to both the number of columns and the number of rows of the matrix. This is reflected more precisely in the following corollary.

Corolario 1. *Given $A \in \mathcal{M}_{m\times n}$ it holds that*

$$\text{rango}(A) \leq m \quad \textit{and} \quad \text{rango}(A) \leq n.$$

Proof. The matrix A has m rows and n columns. By definition, the rank is the size of the largest subset of columns of A that are independent and since in A we have only n columns, the rank will be at most n. In **Theorem 150** we have demonstrated that the rank is also the size of the largest independent set of rows and for that reason again it must be less than the number of rows of the matrix, m. □

We finish by seeing some simple properties of the rank that can simplify some calculations.

Propiedades 2.

i) $\text{rango}\left(\begin{array}{c|c} A & 0 \\ 0 & 0 \end{array}\right) = \text{rango}\begin{pmatrix} A \\ 0 \end{pmatrix} = \text{rango}(A|0) = \text{rango}(A).$

ii) $\text{rango}\left(\begin{array}{c|c} A & B \\ \hline 0 & I_s \end{array}\right) = \text{rango}(A) + s.$

iii) If A_1 is a submatrix of A then $\text{rango}(A_1) \leq \text{rango}(A)$.

Examples 151.

1) Let $B = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & -1 & 2 & 4 \end{pmatrix}$ then it is evident that $\text{rango}(B) \leq 2$ and also, if we consider the submatrix

$$A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$$

and apply elementary operations, we have that

$$A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \xrightarrow[\substack{\text{F1=F1+2F2} \\ \text{-F2}}]{} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and therefore $\text{rango}(A) = 2$ so that, since A is a submatrix of B,

$$2 = \text{rango}(A) \leq \text{rango}(B) \leq 2 \Rightarrow \text{rango}(B) = 2.$$

2) $\text{rango}\begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \\ 1 & 0 & 2 \end{pmatrix} = (\text{F3=F3-F1}) = \text{rango}\begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \\ 0 & -2 & 3 \end{pmatrix} = \begin{pmatrix} \text{C2=C2-2C1} \\ \text{C3=C3+C1} \end{pmatrix} = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & -1 & 2 \\ 0 & -2 & 3 \end{array}\right) =$
$1 + \text{rango}\begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix} = 3.$

Calculations with dependence, independence, linear combinations using the rank

By definition, the rank of a matrix is the largest number of independent columns that we can take in it. It is evident then that if the tuples $v_1, v_2, \ldots, v_m$ are independent, the largest set of independent columns of the matrix $(v_1|v_2|\cdots|v_m)$ will be formed by all those m tuples and consequently we will have

$$\text{rango}(v_1|v_2|\cdots|v_m) = m.$$

In this way, to analyze the dependence or independence of several tuples we can put them in columns and study the rank of the matrix thus obtained. The following property highlights how, following this same idea, we can analyze questions related to dependence, independence and linear combinations through the calculation of ranks.

Property 152. *Consider the n-tuples $v_1, v_2, \ldots, v_m$ and w:*

i) $v_1, v_2, \ldots, v_m$ *are independent* $\Leftrightarrow \text{rango}(v_1|v_2|\cdots|v_m) = m$.

ii) $w \in \langle v_1, v_2, \ldots, v_m\rangle \Leftrightarrow \text{rango}(v_1|v_2|\cdots|v_m) = \text{rango}(v_1|v_2|\cdots|v_m|w)$.

Suppose we have several tuples $v_1, v_2, \ldots, v_m$ and we are working with the set of their linear combinations, $\langle v_1, v_2, \ldots, v_m\rangle$. We know that if the set $\{v_1, v_2, \ldots, v_m\}$ is dependent then some tuples can be obtained as a linear combination of the others and will be superfluous (see **Property 125** on page 140). On the other hand, if they are independent none of them can be eliminated. In the case they are dependent it will be of interest to determine how many and which ones can be eliminated.

If we take into account the definition of rank, **Property 152** also solves the important problem of detecting the superfluous tuples. Indeed, suppose that

$$\text{rango}(v_1|v_2|\cdots|v_m) = r < m.$$

Then, we will be able to find r independent columns within A and, furthermore, no set with more than r columns can be independent. Therefore, it is not difficult to demonstrate that it is enough to find r of those tuples that are independent and we can eliminate all the others.

Examples 153.

1) Check if the tuples $v_1 = (2, 1, 3, -1, 2)$, $v_2 = (4, 2, 6, -2, 4)$, $v_3 = (-2, -1, 0, -5, 1)$, $v_4 = (2, 1, 2, 1, 1)$ and

$v_5 = (2, 1, 5, -5, 4)$ are independent.

We will solve the exercise by applying **Property 152**. To do this, we will begin by calculating the rank of the matrix obtained by putting the tuples in question into columns,

$$\text{rango}(v_1|v_2|v_3|v_4|v_5) = \text{rango}\begin{pmatrix} 2 & 4 & -2 & 2 & 2 \\ 1 & 2 & -1 & 1 & 1 \\ 3 & 6 & 0 & 2 & 5 \\ -1 & -2 & -5 & 1 & -5 \\ 2 & 4 & 1 & 1 & 4 \end{pmatrix} = 2.$$

According to part *i)* of **Property 152)**, p tuples are independent if placed in column they give rank p. However here we have 5 tuples and placed in column they do not give rank 5 but 2. Therefore they are not independent.

2) The tuples from the previous section can be combined to obtain others. Are the five given tuples necessary to obtain all those combinations or, on the contrary, are there superfluous tuples? Let's find a set as small as possible that allows us to obtain all the linear combinations that the initial five tuples provide.

In the previous section we see that, placed in column, the tuples offer rank equal to 2. This indicates to us that the largest independent set will have two tuples. To find that independent set it will be enough to find two tuples independent among the five that we were given at the beginning.

If we take the first two columns we have

$$\text{rango}(v_1|v_2) = \text{rango}\begin{pmatrix} 2 & 4 \\ 1 & 2 \\ 3 & 6 \\ -1 & -2 \\ 2 & 4 \end{pmatrix} = 1$$

and therefore they are not independent (two columns independent would give rank 2) so we can discard this choice. Let's take instead the first and third columns. In this case,

$$\text{rango}(v_1|v_3) = \text{rango}\begin{pmatrix} 2 & -2 \\ 1 & -1 \\ 3 & 0 \\ -1 & -5 \\ 2 & 1 \end{pmatrix} = 2$$

and these two columns are independent. The comments we made after **Property 152** indicate to us that we can keep those two 5-tuples and eliminate the others. That is,

$$\langle v_1, \not{v}_2, v_3, \not{v}_4, \not{v}_5 \rangle = \langle (2,1,3,-1,2), (-2,-1,0,-5,1) \rangle.$$

Therefore, with respect to obtaining linear combinations, if we keep the first and third, all the other tuples are superfluous.

However, instead of trying with the first and third tuples, we could also have chosen, for example, the fourth and the fifth. In that case, placing them in column and calculating their rank we have that

$$\text{rango}(v_4|v_5) = \text{rango}\begin{pmatrix} 2 & 2 \\ 1 & 1 \\ 2 & 5 \\ 1 & -1 \\ 1 & 4 \end{pmatrix} = 2.$$

Therefore, $(2,1,2,1,1)$ and $(2,1,5,-5,4)$ are independent and we can reason as before to conclude that if we keep these two tuples we can eliminate all the others and still obtain the same linear combinations.

In this way, we have two answers for the same exercise. Combining the tuples

$$(2,1,3,-1,2) \quad \text{and} \quad (-2,-1,0,-5,1)$$

we obtain the same combinations as with the five initial tuples and, at the same time, combining

$$(2,1,2,1,1) \quad \text{and} \quad (2,1,5,-5,4)$$

we also obtain all those combinations. Both pairs of tuples represent correct answers to the problem of selecting the simplest set that provides the same combinations as the initial tuples.

3) Check if $w_1 = (1,2,3,2,1)$ and $w_2 = (0,0,3,-6,3)$ can be obtained as a linear combination of the tuples v_1, v_2, v_3, v_4, v_5 from section **1)**.

Let's start with $w_1 = (1,2,3,2,1)$. Applying part *ii)* of **Property 152** we know that

$$w_1 \in \langle v_1, v_2, v_3, v_4, v_5 \rangle \Leftrightarrow \text{rango}(v_1|v_2|v_3|v_4|v_5) = \text{rango}(v_1|v_2|v_3|v_4|v_5|w_1)$$

Therefore, we must check that the rank of the matrix formed by the five initial tuples is equal to that of the matrix obtained by adding the column w_1. That is, we must check if the equality is true

$$¿\,\underbrace{\text{rango}\begin{pmatrix} 2 & 4 & -2 & 2 & 2 \\ 1 & 2 & -1 & 1 & 1 \\ 3 & 6 & 0 & 2 & 5 \\ -1 & -2 & -5 & 1 & -5 \\ 2 & 4 & 1 & 1 & 4 \end{pmatrix}}_{=2 \text{ (section } \mathbf{1)})} = \underbrace{\text{rango}\begin{pmatrix} 2 & 4 & -2 & 2 & 2 & 1 \\ 1 & 2 & -1 & 1 & 1 & 2 \\ 3 & 6 & 0 & 2 & 5 & 3 \\ -1 & -2 & -5 & 1 & -5 & 2 \\ 2 & 4 & 1 & 1 & 4 & 1 \end{pmatrix}}_{=3}?$$

But in section **1)** we saw that the first of the matrices has rank 2 and it is easy to calculate the rank of the second which will be equal to 3. Therefore, the equality of ranks does not occur and as a consequence $w_1 \notin \langle v_1, v_2, v_3, v_4, v_5 \rangle$, that is the tuple $(1, 2, 3, 2, 1)$ cannot be obtained in any way by combining v_1, v_2, v_3, v_4, v_5.

We could have simplified these calculations if we had used the results from section **2)** since then we saw that

$$\langle v_1, v_2, v_3, v_4, v_5 \rangle = \langle v_1, v_3 \rangle.$$

In this way, checking if $w_1 \in \langle v_1, v_2, v_3, v_4, v_5 \rangle$ is equivalent to checking that $w_1 \in \langle v_1, v_3 \rangle$ for which we will have to calculate the rank of matrices of smaller size.

Let's use this last technique for the case of the tuple $w_2 = (0, 0, 3, -6, 3)$. To determine if w_2 can be obtained by combining v_1, v_2, v_3, v_4, v_5, since $\langle v_1, v_2, v_3, v_4, v_5 \rangle = \langle v_1, v_3 \rangle$, it will be enough to study if $w_2 \in \langle v_1, v_3 \rangle$. Now, part *ii)* of **Property 152** will tell us that

$$w_2 \in \langle v_1, v_3 \rangle \Leftrightarrow \text{rango}(v_1|v_3) = \text{rango}(v_1|v_3|w_2).$$

If we calculate the ranks of $(v_1|v_3)$ and $(v_1|v_3|w_2)$ we have

$$\underbrace{\text{rango}\begin{pmatrix} 2 & -2 \\ 1 & -1 \\ 3 & 0 \\ -1 & -5 \\ 2 & 1 \end{pmatrix}}_{=2 \text{ (section } \mathbf{1)})} = \underbrace{\text{rango}\begin{pmatrix} 2 & -2 & 0 \\ 1 & -1 & 0 \\ 3 & 0 & 3 \\ -1 & -5 & -6 \\ 2 & 1 & 3 \end{pmatrix}}_{=2}.$$

Since the ranks coincide, we deduce that indeed $w_2 \in \langle v_1, v_3 \rangle = \langle v_2, v_2, v_3, v_4, v_5 \rangle$ and in this case the tuple $w_2 = (0, 0, 3, -6, 3)$ can be obtained as a linear combination of v_1, v_2, v_3, v_4, v_5.

Column labeling

In the previous examples we have applied **Property 152** to determine which are the superfluous tuples of a given set. Although this technique for finding the tuples independent and superfluous is correct, we will see next that the scheme of elementary operations from page 157 that makes it possible to calculate the rank of a matrix also allows us to detect the independent tuples. To do this it is enough to label each column with an indicator of the tuple to which it corresponds. After applying the scheme of elementary operations, the tuples whose labels appear over the ones on the diagonal of the reduced form will be directly independent.

Example 154. Consider the tuples $v_1 = (1, 2, 0, -1)$, $v_2 = (2, 4, 0, -2)$, $v_3 = (2, 1, 1, 0)$ and $v_4 = (4, 5, 1, -2)$. To see how many and which of them are independent, we will put them in columns and calculate the rank of the resulting matrix. However, we will also label each column with the name we have given to each tuple:

$$\begin{array}{c} \begin{matrix} v_1 & v_2 & v_3 & v_4 \end{matrix} \\ \begin{pmatrix} 1 & 2 & 2 & 4 \\ 2 & 4 & 1 & 5 \\ 0 & 0 & 1 & 1 \\ -1 & -2 & 0 & -2 \end{pmatrix} \end{array} \xrightarrow[\text{C4=C4-C3}]{} \begin{array}{c} \begin{matrix} v_1 & v_2 & v_3 & v_4 \end{matrix} \\ \begin{pmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 1 & 4 \\ 0 & 0 & 1 & 0 \\ -1 & -2 & 0 & -2 \end{pmatrix} \end{array} \xrightarrow[\substack{\text{F1=F1-2F3} \\ \text{F2=F2-F3}}]{} \begin{array}{c} \begin{matrix} v_1 & v_2 & v_3 & v_4 \end{matrix} \\ \begin{pmatrix} 1 & 2 & 0 & 2 \\ 2 & 4 & 0 & 4 \\ 0 & 0 & 1 & 0 \\ -1 & -2 & 0 & -2 \end{pmatrix} \end{array}$$

$$\xrightarrow[\substack{F2=F2-2F1\\ F4=F4+F1}]{} \begin{array}{c} \begin{matrix} v_1 & v_2 & v_3 & v_4 \end{matrix} \\ \begin{pmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \xrightarrow[\substack{C2=C2-2C1\\ C4=C4-2C1}]{} \begin{array}{c} \begin{matrix} v_1 & v_2 & v_3 & v_4 \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \xrightarrow[\substack{C2\leftrightarrow C3\\ F2\leftrightarrow F3}]{} \begin{array}{c} \begin{matrix} v_1 & v_3 & v_2 & v_4 \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Note that if we modify the order of the columns we must modify in the same way the order of the labels. Once we have brought the matrix to its reduced form, it is easy to see that the rank is 2 and therefore we will have at most two tuples independent. But also, the labels that appear over the resulting two ones on the diagonal correspond to the tuples v_1 and v_3 therefore, these two tuples form an independent set that generates the same combinations as the initial four. That is,

$$\langle v_1, v_3 \rangle = \langle v_1, v_2, v_3, v_4 \rangle.$$

In this way, we can discard the tuples v_2 and v_4.

Obtaining all possible tuples

Given a certain set of n-tuples, $v_1, v_2, \ldots, v_m$, in some occasions it is important to determine if they are sufficient to generate via combinations any other n-tuple or if, on the contrary, there are n-tuples that cannot be obtained by combining them.

The set of all n-tuples is $\mathbb{R}^n$, so the question we are posing here is whether with certain tuples $v_1, v_2, \ldots, v_m$ we can obtain all of $\mathbb{R}^n$, that is, whether

$$\langle v_1, v_2, \ldots, v_m \rangle = \mathbb{R}^n.$$

The following property collects results that allow us to address different aspects of this problem.

Property 155.

i) Given the n-tuples $v_1, v_2, \ldots, v_m$*,*

$$\langle v_1, v_2, \ldots, v_m \rangle = \mathbb{R}^n \Leftrightarrow \text{rango}(v_1|v_2|\cdots|v_m) = n.$$

ii) The n-tuple coordinates of $\mathbb{R}^n$*,* $e_1, e_2, \ldots, e_n$ *satisfy*

$$\langle e_1, e_2, \ldots, e_n \rangle = \mathbb{R}^n.$$

iii) Given the n-tuples $v_1, v_2, \ldots, v_n$*,*

$$\langle v_1, v_2, \ldots, v_n \rangle = \mathbb{R}^n \Leftrightarrow \{v_1, v_2, \ldots, v_n\} \text{ is independent.}$$

iv) More than n, n-tuples cannot be independent.

v) With fewer than n, n-tuples it is not possible to obtain all of $\mathbb{R}^n$ *(i.e., all n-tuples).*

Examples 156.
1) Given the 4-tuples,

$$(1, 2, -1, 1),\ (2, 1, 0, -1),\ (0, 0, 1, -1) \text{ and } (1, 2, 0, 1),$$

we ask whether by combining them it is possible to obtain any other 4-tuple we desire.

The set of all 4-tuples is $\mathbb{R}^4$ so we are asking if by combining the given tuples it is possible to obtain all of $\mathbb{R}^4$. Applying part *i)* of **Property 155** we know that

$$\langle (1, 2, -1, 1), (2, 1, 0, -1), (0, 0, 1, -1), (1, 2, 0, 1) \rangle = \mathbb{R}^4$$

will hold if, placed in columns, said tuples give rank 4. We have that

$$\text{rango}\begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 1 & 0 & 2 \\ -1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix} = 4$$

so indeed any tuple of $\mathbb{R}^4$ can be obtained by combining those four. As a consequence of part *iii)* of **Property 155** or of *i)* of **Property 152**, we additionally deduce that the four tuples are independent.

2) Let's study the same problem now referring to the tuples $(1,2,3)$, $(2,1,3)$, $(-1,2,1)$ and $(2,3,5)$. We then ask if by combining them it is possible to obtain any 3-tuple, that is, if

$$\langle (1,2,3),(2,1,3),(-1,2,1),(2,3,5)\rangle = \mathbb{R}^3. \tag{4.4}$$

Without needing to perform any calculation, beforehand, applying part *iv)* of **Property 155** we know that the four tuples are dependent since more than 3 3-tuples are never independent. To check (4.4), as before, with the tuples placed in columns, their rank should be 3. Now,

$$\text{rango}\begin{pmatrix} 1 & 2 & -1 & 2 \\ 2 & 1 & 2 & 3 \\ 3 & 3 & 1 & 5 \end{pmatrix} = 2 < 3$$

so there are tuples of $\mathbb{R}^3$ that cannot be obtained by combining those in this example. For example the tuple $(1,0,0)$ is not a linear combination of them. To check this, we have that

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

It is enough to apply part *ii)* of **Property 152** to deduce that $(1,0,0) \notin \langle (1,2,3),(2,1,3),(-1,2,1),(2,3,5)\rangle$.

4.3.3 Rank and Inverse Matrix. Calculation of the inverse matrix via elementary operations

When we defined the inverse of a matrix in **Section 4.2.4**, we presented a technique for the calculation of inverses (see sections **4)** and **5)** of **Examples 118**) that was valid for the cases of matrices of reduced order. We will use the calculation via elementary operations to offer in this section a new more effective method for handling inverse matrices.

In this section we will study the relationships between the concepts of rank and inverse matrix in the case of square matrices. We will verify that both can be defined one in terms of the other. To begin, we will see in the next property that the fact that a matrix has an inverse, that it is regular, determines its rank directly.

Property 157. *Let $A \in \mathcal{M}_n$ be a square matrix of order n. Then,*

$$\exists A^{-1} \Leftrightarrow \text{rango}(A) = n.$$

Let's make some comments about this last property:

- From the previous property it follows that if a matrix has an inverse, its column tuples must be independent (recall that n n-tuples with rank n are independent). In other words, an equivalent statement for the property would be:

$$\exists A^{-1} \Leftrightarrow \text{the columns of } A \text{ are independent.}$$

On the other hand, if A has an inverse, its transpose matrix, A^t, will also have one, and we can equally deduce that the columns of A^t are independent. Now, the columns of A^t are the rows of A. Therefore, this equivalent statement can also be written in terms of rows.

- As a consequence of **Corollary 1**, a square matrix $A \in \mathcal{M}_n$ of order n (with n rows and n columns) will have rank at most n. Then, n is the maximum value that the rank of A can take and, consequently, it is said that A has maximum rank when rango$(A) = n$. Using this nomenclature, **Property 157** can be rewritten as
$$\exists A^{-1} \Leftrightarrow A \text{ has maximum rank.}$$

Calculation of the inverse matrix via elementary operations

Let's take a square matrix of order n, $A \in \mathcal{M}_n$. If the matrix has rank n then after applying the appropriate elementary operations to it we will arrive at a matrix, of the same size as A, that has all zeros except n ones on the diagonal, that is, we arrive at the identity matrix I_n,

$$A \xrightarrow[\text{elementary operations}]{} I_n.$$

Let's form the matrix $(A|I_n)$ which is obtained by appending the identity matrix to the matrix A. If the elementary operations from before are all by rows and we apply them to $(A|I_n)$, they will transform both the first block (the one corresponding to A) and the second (the one corresponding to I_n). We know that those operations transform A into I_n, $A \to I_n$, and now in addition they will transform I_n into another matrix B, $I_n \to B$. We will therefore have

$$(A|I_n) \xrightarrow[\text{elementary operations}]{} (I_n|B).$$

It can be checked that the matrix B that appears in this way stores the elementary operations that we have applied to A in the sense that when multiplying A by said matrix B we achieve the same effect as when applying the elementary operations to A. That is

$$\underbrace{A \xrightarrow[\text{elementary operations}]{} I_n}_{\text{the el. op. transform } A \text{ into } I_n} \qquad \text{and therefore} \qquad \underbrace{A \cdot B = I_n}_{\text{the product by } B \text{ transforms } A \text{ into } I_n}.$$

But this last equality indicates to us that the matrix B is precisely the inverse matrix of A and therefore $A^{-1} = B$.

The ideas we have outlined in the previous paragraph allow us to determine if a matrix has an inverse and calculate it in that case. In summary, we have the following property:

Property 158. *Suppose that the matrix $A \in \mathcal{M}_n$ has an inverse and that L is a list of elementary operations by rows that transforms A into I_n, then if we apply the operations of L to the block matrix $(A|I_n)$ we will obtain*

$$(A|I_n) \xrightarrow[\text{Operations } L]{} (I_n|B),$$

where B is the inverse of A.

In this way, from **Property 158** it follows that to calculate the inverse of the matrix $A \in \mathcal{M}_n$, we must reduce A to I_n but now applying elementary operations only by rows. For the calculation of the rank we can apply operations both by rows and by columns but now, to obtain the inverse, we have to limit ourselves only to operations by rows. However, this is still possible by applying the Gaussian elimination method that we saw on page 157. We will only have to take into account the following points whose objective is to avoid in all steps the performance of operations by columns:

a) Since we can only apply operations by rows, in step 1 we will only select columns since to cancel the elements of a column the operations to perform are by rows.

b) In step 2, we must take the precaution of selecting an element that is not at the height of the pivots selected in previous steps.

c) We will omit step 4 since it involves performing operations by columns.

d) Once steps 1 to 5 have been applied for all the columns, it will be sufficient to divide the rows in order to transform all the non-zero elements into ones. Finally we will order the rows to obtain the identity.

The procedure from page 157 together with these points allow us to obtain for any regular matrix, A, a list of operations by rows that transform A into I_n. Then, we will use these operations to calculate the inverse as indicated by **Property 158**.

Examples 159.
1) Calculate the inverse of the matrix

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

To do this, we form the matrix $(A \mid I_4)$ and we will apply elementary operations by rows to it until we transform it into $(I_4 \mid B)$.

We will use the matrix reduction method taking into account the points commented before. We start by selecting the third column and its third element as pivot:

$$\left(\begin{array}{cccc|cccc} 1 & 2 & \mathbf{0} & 1 & 1 & 0 & 0 & 0 \\ 1 & -1 & \mathbf{2} & 1 & 0 & 1 & 0 & 0 \\ 2 & 0 & \underline{\mathbf{1}} & 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & \mathbf{2} & 3 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[\substack{\text{F2=F2-2F3} \\ \text{F4=F4-2F3}}]{} \left(\begin{array}{cccc|cccc} 1 & 2 & \mathbf{0} & 1 & 1 & 0 & 0 & 0 \\ -3 & -1 & \mathbf{0} & 1 & 0 & 1 & -2 & 0 \\ 2 & 0 & \underline{\mathbf{1}} & 0 & 0 & 0 & 1 & 0 \\ -3 & -1 & \mathbf{0} & 3 & 0 & 0 & -2 & 1 \end{array}\right)$$

As we have said we cannot perform operations by columns and for that reason we cannot cancel the row perpendicular at the height of the pivot. Therefore, we select a new column and a new pivot that, as already mentioned before, should not be at the height of the other chosen pivots (so far only the **1** from the previous step). Let's now select the fourth column and take the second element as pivot

$$\left(\begin{array}{cccc|cccc} 1 & 2 & 0 & \mathbf{1} & 1 & 0 & 0 & 0 \\ -3 & -1 & 0 & \underline{\mathbf{1}} & 0 & 1 & -2 & 0 \\ 2 & 0 & \underline{1} & \mathbf{0} & 0 & 0 & 1 & 0 \\ -3 & -1 & 0 & \mathbf{3} & 0 & 0 & -2 & 1 \end{array}\right) \xrightarrow[\substack{\text{F1=F1-F2} \\ \text{F4=F4-3F2}}]{} \left(\begin{array}{cccc|cccc} 4 & 3 & 0 & \mathbf{0} & 1 & -1 & 2 & 0 \\ -3 & -1 & 0 & \underline{\mathbf{1}} & 0 & 1 & -2 & 0 \\ 2 & 0 & \underline{1} & \mathbf{0} & 0 & 0 & 1 & 0 \\ 6 & 2 & 0 & \mathbf{0} & 0 & -3 & 4 & 1 \end{array}\right).$$

We continue selecting the second column and taking as pivot the fourth element to avoid it being at the height of the two previous ones. Before canceling the other elements, to simplify the calculations, we convert the pivot to 1 by dividing the fourth row by 2:

$$\xrightarrow[\frac{1}{2}\text{F4}]{} \left(\begin{array}{cccc|cccc} 4 & \mathbf{3} & 0 & 0 & 1 & -1 & 2 & 0 \\ -3 & \mathbf{-1} & 0 & \underline{1} & 0 & 1 & -2 & 0 \\ 2 & \mathbf{0} & \underline{1} & 0 & 0 & 0 & 1 & 0 \\ 3 & \underline{\mathbf{1}} & 0 & 0 & 0 & \frac{-3}{2} & 2 & \frac{1}{2} \end{array}\right) \xrightarrow[\substack{\text{F1=F1-3F4} \\ \text{F2=F2+F4}}]{} \left(\begin{array}{cccc|cccc} -5 & \mathbf{0} & 0 & 0 & 1 & \frac{7}{2} & -4 & \frac{-3}{2} \\ 0 & \mathbf{0} & 0 & \underline{1} & 0 & \frac{-1}{2} & 0 & \frac{1}{2} \\ 2 & \mathbf{0} & \underline{1} & 0 & 0 & 0 & 1 & 0 \\ 3 & \underline{\mathbf{1}} & 0 & 0 & 0 & \frac{-3}{2} & 2 & \frac{1}{2} \end{array}\right).$$

We finish by selecting the first column and as pivot, if we intend it not to be at the height of those selected so far, we can only choose the first one. Given that the value of the pivot is -5 we first divide the first row to transform it into 1:

$$\xrightarrow[\frac{-1}{5}\text{F1}]{} \left(\begin{array}{cccc|cccc} \underline{\mathbf{1}} & 0 & 0 & 0 & \frac{-1}{5} & \frac{-7}{10} & \frac{4}{5} & \frac{3}{10} \\ \mathbf{0} & 0 & 0 & \underline{1} & 0 & \frac{-1}{2} & 0 & \frac{1}{2} \\ \mathbf{2} & 0 & \underline{1} & 0 & 0 & 0 & 1 & 0 \\ \mathbf{3} & \underline{1} & 0 & 0 & 0 & \frac{-3}{2} & 2 & \frac{1}{2} \end{array}\right) \xrightarrow[\substack{\text{F3=F3-2F1} \\ \text{F4=F4-3F1}}]{} \left(\begin{array}{cccc|cccc} \underline{\mathbf{1}} & 0 & 0 & 0 & \frac{-1}{5} & \frac{-7}{10} & \frac{4}{5} & \frac{3}{10} \\ \mathbf{0} & 0 & 0 & \underline{1} & 0 & \frac{-1}{2} & 0 & \frac{1}{2} \\ \mathbf{0} & 0 & \underline{1} & 0 & \frac{2}{5} & \frac{7}{5} & \frac{-3}{5} & \frac{-3}{5} \\ \mathbf{0} & \underline{1} & 0 & 0 & \frac{3}{5} & \frac{3}{5} & \frac{-2}{5} & \frac{-2}{5} \end{array}\right).$$

To finish it is enough to reorder

$$\xrightarrow[\text{reordering}]{} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & \frac{-1}{5} & \frac{-7}{10} & \frac{4}{5} & \frac{3}{10} \\ 0 & 1 & 0 & 0 & \frac{3}{5} & \frac{3}{5} & \frac{-2}{5} & \frac{-2}{5} \\ 0 & 0 & 1 & 0 & \frac{2}{5} & \frac{7}{5} & \frac{-3}{5} & \frac{-3}{5} \\ 0 & 0 & 0 & 1 & 0 & \frac{-1}{2} & 0 & \frac{1}{2} \end{array}\right)$$

from which we obtain,

$$A^{-1} = \begin{pmatrix} \frac{-1}{5} & \frac{-7}{10} & \frac{4}{5} & \frac{3}{10} \\ \frac{3}{5} & \frac{3}{5} & \frac{-2}{5} & \frac{-2}{5} \\ \frac{2}{5} & \frac{7}{5} & \frac{-3}{5} & \frac{-3}{5} \\ 0 & \frac{-1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

2) Consider $A = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ and let's see that it is a regular matrix:

$$\begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \xrightarrow[\text{C3=C3-C1}]{} \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} \xrightarrow[\text{C3=C3+C2}]{} \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow[\frac{1}{2}\text{F1}]{} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \xrightarrow[\text{ordering columns}]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3.$$

Therefore rango$(A) = 3$ and consequently the matrix A is regular. We can now calculate its inverse by applying elementary operations to the rows of the matrix $\left(\begin{array}{c|c} A & I_3 \end{array} \right)$

$$\left(\begin{array}{c|c} A & I_3 \end{array} \right) = \left(\begin{array}{ccc|ccc} 0 & 0 & 2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[\frac{1}{2}\text{F1}]{} \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[\text{F2=F2-F1}]{} \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & -\frac{1}{2} & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[\text{F3=F3+F1}]{} \left(\begin{array}{ccc|ccc} 0 & 0 & 1 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & -\frac{1}{2} & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & 1 \end{array}\right)$$

$$\xrightarrow[\text{reordering rows}]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & 0 \end{array}\right) = \left(\begin{array}{c|c} I_3 & B \end{array} \right).$$

Therefore:

$$A^{-1} = B = \begin{pmatrix} -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}.$$

3) Applying elementary transformations to $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ we achieve:

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \xrightarrow[\text{F2=F2-F1}]{} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \xrightarrow[\text{C2=C2-C1}]{} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Therefore rango$(A) = 1 < 2$ and consequently A is a singular matrix that does not have an inverse.

4) *Calculate the matrix C that satisfies the following equality:*

$$\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \cdot C = \begin{pmatrix} 1 \\ 6 \end{pmatrix}.$$

The simplest way to calculate C is to solve for the matrix $\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$. To do this, first we must check if it is regular:

$$\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \xrightarrow[\text{F2=F2+F1}]{} \begin{pmatrix} 1 & 2 \\ 0 & 5 \end{pmatrix} \xrightarrow[\text{C2=C2-2C1}]{} \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$$
$$\xrightarrow[\frac{1}{5}\text{F2}]{} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

Through the chain of operations above we obtain I_2 so the matrix in question has rank 2 and is indeed regular so we can solve for it:

$$C = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 \\ 6 \end{pmatrix}.$$

We now calculate the inverse of the matrix:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{array}\right) \xrightarrow[\text{F2=F2+F1}]{} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 5 & 1 & 1 \end{array}\right) \xrightarrow[\frac{1}{5}\text{F2}]{} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{1}{5} & \frac{1}{5} \end{array}\right)$$
$$\xrightarrow[\text{F1=F1-2F2}]{} \left(\begin{array}{cc|cc} 1 & 0 & \frac{3}{5} & -\frac{2}{5} \\ 0 & 1 & \frac{1}{5} & \frac{1}{5} \end{array}\right),$$

so finally

$$\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}^{-1} = \frac{1}{5} \cdot \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix},$$

and therefore:

$$C = \frac{1}{5} \cdot \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 6 \end{pmatrix} = \frac{1}{5} \cdot \begin{pmatrix} -9 \\ 7 \end{pmatrix} = \begin{pmatrix} -9/5 \\ 7/5 \end{pmatrix}.$$

To finish this section, we will give an alternative definition of rank based on the concept of inverse. Given $A \in \mathcal{M}_n$ we have seen that

$$\text{rango}(A) = n \Leftrightarrow \exists A^{-1}$$

However, if $\text{rango}(A) < n$ or A is not square, this characterization does not make sense. For an arbitrary matrix, not necessarily square, $A \in \mathcal{M}_{m\times n}$, is it possible to characterize the fact that the rank reaches a specific value, $\text{rango}(A) = r$, in a similar way?

The following property answers this question. We will see that if $\text{rango}(A) = r$ we probably will not be able to calculate the inverse of A but there will exist within A a square submatrix of order r that will have an inverse. We handle here the concept of minor that we define next:

Definition 160. Given $A \in \mathcal{M}_{m\times n}$, we call a minor of order r of the matrix A any square submatrix of order r of A.

Example 161. Let $A = \begin{pmatrix} 2 & 1 & 3 & 6 \\ -1 & 2 & 2 & 4 \\ 0 & 0 & 1 & 2 \end{pmatrix}$ we have that:

- ⋆ $\begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ is a square submatrix of order 3 of A and therefore it is a minor of order 3 of the matrix A.
- ⋆ $\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$ is a square submatrix of order 2 of A and for that reason it is a minor of order 2 of A.
- ⋆ $\begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}$ is a square submatrix of order 2 of A and as a consequence a minor of order 2.

⋆ (2), (0) or (4) are square submatrices of order 1, that is, minors of order 1.

Let's now see the announced property. Note that the characterization we give in it is, in its statement, similar to the definition we gave for the rank in **Definition 138** (pag. 152).

Property 162. *Given $A \in \mathcal{M}_{m\times n}$,*

$$\text{rango}(A) = r \Leftrightarrow \textit{the order of the largest regular minor of } A \textit{ is } r.$$

In other words, if a matrix has rank r, necessarily we will be able to find within it a square submatrix, a minor, of size r with inverse and furthermore it is not possible to find larger submatrices with inverse.

Aquí está la traducción al inglés del fragmento sobre determinantes, cumpliendo estrictamente con todas las indicaciones.

4.4 Determinant of a Matrix

We have seen how to find out if a square matrix is regular or not. We will study in this section alternative mechanisms to determine regularity as well as for the calculation of the inverse.

The tool we will use for this purpose is the determinant. The determinant is a real number that, in a certain sense, measures the degree to which a set of n n-tuples are independent. If the determinant is zero the n-tuples will be dependent and if it is different from zero they will be independent but in such a way that the closer the determinant is to zero the more close, in a certain sense, to being dependent the n-tuples will be.

For the definition of the determinant we will follow a constructive method based on the formulas for the expansion of the determinant by a row or column.

Definition 163. Given $A \in \mathcal{M}_n$ we define the determinant of A and denote it by $\det(A)$ or $|A|$, as the number that satisfies:

- If $A = (a)_{1\times 1} \in \mathcal{M}_1$, then $|A| = |(a)| = a$.
- If $A = (a_{ij})_{n\times n} \in \mathcal{M}_n$ with $n > 1$, then $|A|$ is defined in either of the following two ways:

  a) For any $i = 1, \ldots, n$:

$$\begin{aligned}|A| &= a_{i1}\cdot\Delta_{i1} + a_{i2}\cdot\Delta_{i2} + \cdots + a_{in}\cdot\Delta_{in} \\ &= \underbrace{\begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{in}\end{pmatrix}}_{\text{row } i \text{ of } A}\cdot\begin{pmatrix}\Delta_{i1}\\ \Delta_{i2}\\ \vdots \\ \Delta_{in}\end{pmatrix}.\end{aligned}$$

  The above formula is known as the expansion of the determinant of the matrix A along the i-th row.

  b) For any $j = 1, \ldots, n$:

$$\begin{aligned}|A| &= a_{1j}\cdot\Delta_{1j} + a_{2j}\cdot\Delta_{2j} + \cdots + a_{nj}\cdot\Delta_{nj} \\ &= \begin{pmatrix}\Delta_{1j} & \Delta_{2j} & \ldots & \Delta_{nj}\end{pmatrix}\cdot\underbrace{\begin{pmatrix} a_{1j}\\ a_{2j}\\ \vdots \\ a_{nj}\end{pmatrix}}_{\text{column } j \text{ of } A}.\end{aligned}$$

  This formula is called the expansion of the determinant of the matrix A along the j-th column.

Where Δ_{ij} is called the (i, j) cofactor of the matrix and is defined by the formula

$$\Delta_{ij} = (-1)^{i+j} \cdot |A_{(i,j)}|,$$

where $A_{(i,j)}$ is the submatrix of A obtained by deleting the i-th row and the j-th column.

Remark. Although the symbol we use for the determinant of a matrix, $|\cdot|$, is the same as for the absolute value of a real number, they are totally different concepts and have no relation.

Examples 164.
1) To calculate the determinant of the matrix $(3)_{1\times1}$ we will refer to the first point of the definition of determinant since this matrix is a 1×1 type matrix. Then we have:

$$|(3)| = 3.$$

$|(-5)| = -5$ (remember that $|(-5)|$ is the determinant of the matrix $(-5)_{1\times1}$ and not the absolute value of the number -5).

2) *Determinant of the square matrix of order 2.*

Consider $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathcal{M}_2$ and let's try to calculate its determinant. Since it is not a 1×1 matrix we cannot apply the first point of the definition of determinant and we must refer to the second. For this we have to:

* $A_{(1,1)}$ is the submatrix of A that we obtain by deleting row 1 and column 1:

$$A_{(1,1)} = (a_{22}).$$

* $A_{(1,2)}$ is the submatrix of A that we obtain by deleting row 1 and column 2:

$$A_{(1,2)} = (a_{21}).$$

* $A_{(2,1)}$ is the submatrix of A that we obtain by deleting row 2 and column 1:

$$A_{(2,1)} = (a_{12}).$$

* $A_{(2,2)}$ is the submatrix of A that we obtain by deleting row 2 and column 2:

$$A_{(2,2)} = (a_{11}).$$

From the above we calculate the cofactors of the matrix:

$$\begin{aligned} \Delta_{11} &= (-1)^{1+1} \cdot |A_{(1,1)}| = (-1)^2 \cdot |(a_{22})| = a_{22}, \\ \Delta_{12} &= (-1)^{1+2} \cdot |A_{(1,2)}| = (-1)^3 \cdot |(a_{21})| = -a_{21}, \\ \Delta_{21} &= (-1)^{2+1} \cdot |A_{(2,1)}| = (-1)^3 \cdot |(a_{12})| = -a_{12}, \\ \Delta_{22} &= (-1)^{2+2} \cdot |A_{(2,2)}| = (-1)^4 \cdot |(a_{11})| = a_{11}. \end{aligned}$$

For the calculation of the determinant we now have four possibilities:

1. Do the expansion of the determinant along the first row.
2. Do the expansion of the determinant along the second row.
3. Do the expansion of the determinant along the first column.

4. Do the expansion of the determinant along the second column.

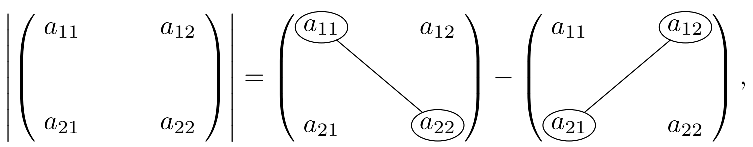
We will calculate following two of the previous possibilities:

$$\begin{aligned}
\left|\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right| &= \\
\text{(row 1)} &= a_{11} \cdot \Delta_{11} + a_{12} \cdot \Delta_{12} = a_{11} \cdot a_{22} + a_{12} \cdot (-a_{21}) \\
&= a_{11} \cdot a_{22} - a_{12} \cdot a_{21}. \\
\text{(column 2)} &= a_{12} \cdot \Delta_{12} + a_{22} \cdot \Delta_{22} = a_{12} \cdot (-a_{21}) + a_{22} \cdot a_{11} \\
&= a_{11} \cdot a_{22} - a_{12} \cdot a_{21}.
\end{aligned}$$

Regardless of the method followed (expansion along the first row or along the second column) we have obtained the same value for the determinant. It can be checked as an exercise that if we expand along the second row or along the first column the result will also be the same. In summary we have obtained the following formula for the determinant of a 2×2 matrix:

$$\left|\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right| = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}.$$

Schematically, we can represent this formula as follows



where the elements that appear joined by a segment must be multiplied.

3) *Determinant of a square matrix of order 3.*

Given the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

we will calculate its determinant by expanding along the first column. We will then have:

$$|A| = a_{11} \cdot \Delta_{11} + a_{21} \cdot \Delta_{21} + a_{31} \cdot \Delta_{31}.$$

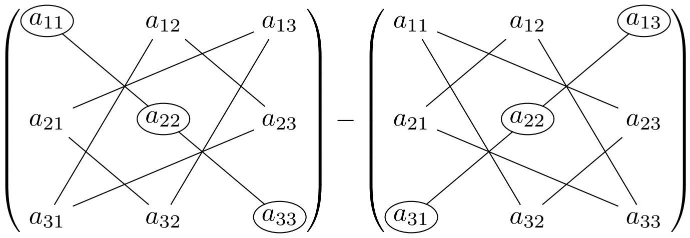
We next calculate the necessary cofactors in the above formula:

$$\begin{aligned}
\Delta_{11} &= (-1)^{1+1}|A_{(1,1)}| = \left|\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}\right| = a_{22} \cdot a_{33} - a_{23} \cdot a_{32}, \\
\Delta_{21} &= (-1)^{2+1}|A_{(2,1)}| = -\left|\begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix}\right| = -a_{12} \cdot a_{33} + a_{32} \cdot a_{13}, \\
\Delta_{31} &= (-1)^{3+1}|A_{(3,1)}| = \left|\begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}\right| = a_{12} \cdot a_{23} - a_{22} \cdot a_{13}.
\end{aligned}$$

We then have:

$$\begin{aligned}
|A| &= a_{11} \cdot (a_{22} \cdot a_{33} - a_{23} \cdot a_{32}) + a_{21} \cdot (a_{32} \cdot a_{13} - a_{12} \cdot a_{33}) \\
&\quad + a_{31} \cdot (a_{12} \cdot a_{23} - a_{22} \cdot a_{13}). \\
&= a_{11} \cdot a_{22} \cdot a_{33} + a_{21} \cdot a_{32} \cdot a_{13} + a_{31} \cdot a_{12} \cdot a_{23} \\
&\quad - (a_{31} \cdot a_{22} \cdot a_{13} + a_{11} \cdot a_{23} \cdot a_{32} + a_{21} \cdot a_{12} \cdot a_{33}).
\end{aligned}$$

The above formula for the determinant of a 3×3 matrix is called Sarrus' rule and can be schematized in the following diagram:



4) Next we calculate some determinants of 2×2 matrices using the formula obtained in example 3:

$\star \left|\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}\right| = 1 \cdot 0 - 2 \cdot (-1) = 2.$

$\star \left|\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right| = 1 \cdot 1 - 1 \cdot 1 = 0.$

$\star \left|\begin{pmatrix} 1 & 6 \\ 3 & 9 \end{pmatrix}\right| = 1 \cdot 9 - 6 \cdot 3 = -9.$

5) We now use Sarrus' formula to calculate the determinant of 3×3 matrices:

$$\begin{aligned} \star \left|\begin{pmatrix} 1 & 6 & 0 \\ -1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}\right| &= \\ &= 1 \cdot 2 \cdot 1 + (-1) \cdot 0 \cdot 0 + 6 \cdot 0 \cdot 1 - (0 \cdot 2 \cdot 1 + 0 \cdot 0 \cdot 1 + (-1) \cdot 6 \cdot 1) \\ &= 2 + 6 = 8. \end{aligned}$$

$$\begin{aligned} \star \left|\begin{pmatrix} 4 & 2 & 6 \\ 0 & 3 & -1 \\ 0 & 0 & 5 \end{pmatrix}\right| &= \\ &= 4 \cdot 3 \cdot 5 + 0 \cdot 0 \cdot 6 + 2 \cdot (-1) \cdot 0 - (0 \cdot 3 \cdot 6 + 0 \cdot (-1) \cdot 4 + 0 \cdot 2 \cdot 5) \\ &= 4 \cdot 3 \cdot 5 = 60. \end{aligned}$$

6) We will now calculate the determinant of a 3×3 matrix by expanding along the second row instead of using Sarrus' formula:

$$\begin{aligned} \left|\begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix}\right| &= \\ &= 0 \cdot \Delta_{21} + 1 \cdot \Delta_{22} + (-1) \cdot \Delta_{23} \\ &= 1 \cdot (-1)^{2+2} \left|\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}\right| + (-1) \cdot (-1)^{2+3} \left|\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}\right| \\ &= (1 \cdot 0 - (-1) \cdot 0) + (1 \cdot (-1) - 0 \cdot 2) = -1. \end{aligned}$$

7) We next solve the determinant of a 4×4 matrix. To do this we have only to apply the expansion formulas by rows or by columns. In general the most suitable row or column to expand along will be the one that has a greater number of zeros since then the amount of calculations to be performed will be reduced. In this case we will do an expansion along the third column:

$$\begin{aligned}
\left|\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 3 & 0 & 2 \\ 1 & 1 & -1 & 0 \end{pmatrix}\right| &= 0 \cdot \Delta_{13} + 1 \cdot \Delta_{23} + 0 \cdot \Delta_{33} - \Delta_{43} \\
&= (-1)^{2+3} \left|\begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \\ 1 & 1 & 0 \end{pmatrix}\right| - (-1)^{4+3} \left|\begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \\ 0 & 3 & 2 \end{pmatrix}\right| \\
&= -(0+0+4-0-3-2) + (-2+0+0-0-0-0) \\
&= 1 - 2 = -1.
\end{aligned}$$

8) In the following example we solve a determinant of type 4×4 by an expansion along the second row:

$$\begin{aligned}
&\left|\begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & 1 & 2 & 2 \end{pmatrix}\right| = \\
&= 2 \cdot \Delta_{21} + 0 \cdot \Delta_{22} + \Delta_{23} + \Delta_{24} \\
&= -2 \cdot \left|\begin{pmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 2 \end{pmatrix}\right| - \left|\begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}\right| + \left|\begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}\right| \\
&= -2 \cdot (4+2-1-1-4+2) - (2+1-2+1-1-4) \\
&\quad +(2-1-2-1-1-4) = -4+3-7 = -8.
\end{aligned}$$

In the previous examples it becomes clear that the determinant can take any real number as its value. On the other hand, the rank will always be a natural number.

The following result collects fundamental properties of the determinant. Among them, those that highlight the relationship of the determinant with the product and regularity of matrices are important.

Propiedades 3.

1. *Given $A = (a_{ij})_{n\times n} \in \mathcal{M}_n$ a diagonal, lower triangular or upper triangular matrix, it holds that:*

$$|A| = a_{11} \cdot a_{22} \cdot a_{33} \cdot \cdots \cdot a_{nn}.$$

   In particular $|I_n| = 1$.

2. *Given $A \in \mathcal{M}_n$, $|A^t| = |A|$.*
3. *Given $A, B \in \mathcal{M}_n$, $|A \cdot B| = |A| \cdot |B|$.*
4. *Given $A \in \mathcal{M}_n$ regular, we have that $|A| \neq 0$ and also*

$$|A^{-1}| = \frac{1}{|A|}.$$

5. *Given $A \in \mathcal{M}_n$ such that $|A| \neq 0$ it holds that A is regular.*
6. *Given a matrix $B \in \mathcal{M}_{m\times n}$,*

$$\text{rango}(B) = r \Leftrightarrow \textit{The order of the largest minor of } A \textit{ with non-zero determinant is } r.$$

Proof. We will see the proof only of properties 1 and 4:

1. We will only see the case of upper triangular matrices. For the rest of the cases the proof is similar.

   Let $A \in \mathcal{M}_n$ be an upper triangular matrix of the form:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

   Then we have:

$$\begin{aligned} |A| &= \begin{pmatrix}\text{expanding along} \\ \text{column 1}\end{pmatrix} = a_{11} \cdot (-1)^{1+1} \cdot \left| \begin{pmatrix} a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ 0 & a_{33} & a_{34} & \cdots & a_{3n} \\ 0 & 0 & a_{44} & \cdots & a_{4n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix} \right| \\ &= \begin{pmatrix}\text{expanding along the} \\ \text{new column 1}\end{pmatrix} = a_{11} \cdot a_{22} \cdot (-1)^{1+1} \cdot \left| \begin{pmatrix} a_{33} & a_{34} & a_{35} & \cdots & a_{3n} \\ 0 & a_{44} & a_{45} & \cdots & a_{4n} \\ 0 & 0 & a_{55} & \cdots & a_{5n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix} \right| \\ &= \begin{pmatrix}\text{repeating the process} \\ \text{successively}\end{pmatrix} = a_{11} \cdot a_{22} \cdot \cdots \cdot a_{nn}. \end{aligned}$$

4. If A is regular then its inverse matrix A^{-1} will exist and will satisfy:

$$A \cdot A^{-1} = I_n.$$

   Taking the determinant on both sides of this equality and applying property 3 we obtain:

$$|A \cdot A^{-1}| = |I_n| \Rightarrow |A| \cdot |A^{-1}| = 1$$

$$\Rightarrow \begin{cases} |A| \neq 0 \\ |A^{-1}| = \frac{1}{|A|} \text{ (solving for } |A^{-1}| \text{ in the previous equality)} \end{cases}.$$

□

Remark. Properties 4 and 5 are important since they indicate that we can use the determinant to check whether a matrix is regular or not by simply checking whether the determinant is zero or not.

Examples 165.

1) Considering the matrix $A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$ we have:

$$|A| = 1 \cdot 3 - 2 \cdot (-1) = 5 \neq 0.$$

Since the determinant of A is non-zero we know that it is a regular matrix and also without needing to know its inverse, A^{-1}, we know the value of its determinant:

$$|A^{-1}| = \frac{1}{|A|} = \frac{1}{5}.$$

2) $\left|\begin{pmatrix} 2 & 1 & 3 \\ 0 & -3 & 4 \\ 0 & 0 & 6 \end{pmatrix}\right| = \left(\begin{matrix}\text{using} \\ \text{property 1}\end{matrix}\right) = 2 \cdot (-3) \cdot 6 = -36.$

3) Part *6* of **Property 3** provides an alternative technique for the calculation of the rank. We can determine the rank of a matrix by applying elementary operations or through this last property.

We will calculate the rank of the matrix $A = \begin{pmatrix} 2 & 1 & 3 & 6 \\ -1 & 2 & 2 & 4 \\ 0 & 0 & 1 & 2 \end{pmatrix}$ using the two techniques we know:

i) The largest square submatrix of A will be of order at most 3 so the rank of the matrix will have a maximum value of 3. To check if the rank is 3 we must find a submatrix 3×3 with non-zero determinant. But this is easy since if we take the submatrix formed by the first three columns of A we have

$$\left|\begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}\right| = 5 \neq 0,$$

so we have a minor of order 3 that is non-zero and the rank will be 3 since there cannot be larger minors in the matrix A. So applying this definition we obtain:

$$\text{rango}(A) = 3.$$

ii) We will apply elementary operations to the matrix A until transforming it into a matrix with ones on the diagonal:

$$A \xrightarrow[\text{C4=C4-2C3}]{} \begin{pmatrix} 2 & 1 & 3 & 0 \\ -1 & 2 & 2 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \xrightarrow[\text{F1=F1-3F3}]{\text{F2=F2-2F3}} \begin{pmatrix} 2 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow[\text{C2=C2+2C1}]{} \begin{pmatrix} 2 & 5 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \xrightarrow[\text{F1=F1+2F2}]{} \begin{pmatrix} 0 & 5 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow[\text{-F2}]{\frac{1}{5}\text{F1}} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \xrightarrow[\text{ordering}]{} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

We have obtained a matrix with three ones on the diagonal and zero in all other positions so finally:

$$\text{rango}(A) = 3.$$

4) Let $A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 3 & 1 \\ 1 & 4 & 5 & 2 \end{pmatrix}$ and let's calculate its rank using the two methods we know:

i) The largest square matrix that can be found within A is of type 3×3 so the largest possible rank for A is 3. Let's see if A has any non-zero minor of order 3 and for this let's consider all possible square submatrices of order 3 of A and obtain their determinant:

$$\left\{\begin{array}{ll} \left|\begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ 1 & 4 & 5 \end{pmatrix}\right| = 0 & \left|\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 4 & 2 \end{pmatrix}\right| = 0 \\ \left|\begin{pmatrix} 1 & -1 & 0 \\ 0 & 3 & 1 \\ 1 & 5 & 2 \end{pmatrix}\right| = 0 & \left|\begin{pmatrix} 2 & -1 & 0 \\ 1 & 3 & 1 \\ 4 & 5 & 2 \end{pmatrix}\right| = 0 \end{array}\right. .$$

The above are all possible 3×3 submatrices and all have zero determinant so the rank of the matrix will not be 3. However, it will have rank 2 since it is easy to find submatrices of A with non-zero determinant such as for example:

$$\left|\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\right| = 1 \neq 0.$$

In this way the order of the largest minor with non-zero determinant is 2 and therefore the rank of the matrix will be:

$$\text{rango}(A) = 2.$$

ii) Performing operations on A we have:

$$A \xrightarrow[\text{C3=C3+C1}]{} \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 3 & 1 \\ 1 & 4 & 6 & 2 \end{pmatrix} \xrightarrow[\text{C2=C2-2C1}]{} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 \\ 1 & 2 & 6 & 2 \end{pmatrix}$$

$$\xrightarrow[\text{F3=F3-F1}]{} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 2 & 6 & 0 \end{pmatrix} \xrightarrow[\text{C3=C3-3C2}]{} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow[\text{F3=F3-2F2}]{} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Therefore

$$\text{rango}(A) = \text{rango}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 2.$$

Propiedades 4 (Elementary operations for determinants)**.** *Given $A = (a_{ij})_{n\times n} \in \mathcal{M}_n$ we have that:*

1. *A number that multiplies an entire row or an entire column can be taken out of the determinant. That is, $\forall i, j \in \{1, \ldots, n\}$*

$$\left|\begin{pmatrix} a_{11} & \cdots & r\cdot a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & r\cdot a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & r\cdot a_{nj} & \cdots & a_{nn} \end{pmatrix}\right| = r\cdot\left|\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}\right|,$$

$$\left|\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ r\cdot a_{i1} & r\cdot a_{i2} & \cdots & r\cdot a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\right| = r\cdot\left|\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\right|.$$

2. *If we interchange one column (respectively row) with another column (resp. row) that is adjacent, the determinant changes sign.*

3. *If to one column (respectively row) we add another column (resp. row) multiplied by a number, the determinant does not change.*

Examples 166.

1. Applying property 1 we have:

   ⋆ $\left|\begin{pmatrix}1 & 2 & 0\\ 6 & 4 & 3\\ 2 & 8 & 1\end{pmatrix}\right| = \left|\begin{pmatrix}1 & 2\cdot 1 & 0\\ 6 & 2\cdot 2 & 3\\ 2 & 2\cdot 4 & 1\end{pmatrix}\right| = 2\cdot\left|\begin{pmatrix}1 & 1 & 0\\ 6 & 2 & 3\\ 2 & 4 & 1\end{pmatrix}\right|.$

   ⋆ $\left|\begin{pmatrix}1 & 2 & 3\\ 0 & \frac{2}{5} & \frac{4}{5}\\ 2 & -9 & 1\end{pmatrix}\right| = \left|\begin{pmatrix}1 & 2 & 3\\ \frac{1}{5}\cdot 0 & \frac{1}{5}\cdot 2 & \frac{1}{5}\cdot 4\\ 2 & -9 & 1\end{pmatrix}\right| = \frac{1}{5}\cdot\left|\begin{pmatrix}1 & 2 & 3\\ 0 & 2 & 4\\ 2 & -9 & 1\end{pmatrix}\right|.$

2. Applying property 2 repeatedly we have:

$$\begin{aligned}\left|\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\end{pmatrix}\right| &= (\text{C2} \leftrightarrow \text{C3}) = -\left|\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}\right|\\ &= (\text{C3} \leftrightarrow \text{C4}) = \left|\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\right|\\ &= (\text{C2} \leftrightarrow \text{C3}) = -\left|\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\right| = -|I_n| = -1.\end{aligned}$$

3. In the following case we will apply property 3 repeatedly to simplify the calculation of the determinant:

$$\begin{aligned}&\left|\begin{pmatrix}1 & 1 & -3\\ -1 & 6 & 3\\ 0 & 4 & 4\end{pmatrix}\right| =\\ &= (\text{C3=C3+3C1}) = \left|\begin{pmatrix}1 & 1 & 0\\ -1 & 6 & 0\\ 0 & 4 & 4\end{pmatrix}\right|\\ &= (\text{F2=F2+F1}) = \left|\begin{pmatrix}1 & 1 & 0\\ 0 & 7 & 0\\ 0 & 4 & 4\end{pmatrix}\right|\\ &= (\text{C2=C2-C3}) = \left|\begin{pmatrix}1 & 1 & 0\\ 0 & 7 & 0\\ 0 & 0 & 4\end{pmatrix}\right|\\ &= 1\cdot 7\cdot 4 = 28.\end{aligned}$$

The last property together with the others we have seen for determinants allow us to solve some simple determinants directly. Let's see it in the following note:

Remark. It is important to keep in mind the following points:

- If $A \in \mathcal{M}_n$ has an entire column of zeros or an entire row of zeros then $|A| = 0$. This is evident since if all the elements in the row or column are zero, expanding along that row or column we will obtain zero as a result.
- If $A \in \mathcal{M}_n$ has two equal columns or two equal rows then $|A| = 0$. This is clear since in such a case applying part 3 of **Properties 4**, we can subtract from the row or column the other one that coincides with it, thus obtaining a row or column of zeros so the determinant will be zero.

Example 167. $\left|\begin{pmatrix} 4 & -3 & 5 \\ -1 & -1 & 0 \\ 4 & -3 & 5 \end{pmatrix}\right| = \text{(F3=F3-F1)} = \left|\begin{pmatrix} 4 & -3 & 5 \\ -1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right| = 0.$

- If some column (respectively row) can be obtained by adding the other columns (resp. rows) multiplied by numbers, then $|A| = 0$. Indeed if a column is obtained by adding in a certain way the other columns of the matrix, we can subtract appropriately from this column until we obtain a complete column of zeros (we can make the same comment for rows). In other words, if the rows or columns of the matrix are dependent then the determinant will be zero.

Example 168. In the matrix of the following determinant the third column can be obtained as the sum of twice the first column plus three times the second, so:

$$\left|\begin{pmatrix} 1 & 1 & 5 \\ -1 & 6 & 16 \\ 0 & 4 & 12 \end{pmatrix}\right| = \text{(C3=C3-2C1)} = \left|\begin{pmatrix} 1 & 1 & 3 \\ -1 & 6 & 18 \\ 0 & 4 & 12 \end{pmatrix}\right| = \text{(C3=C3-3C2)} = \left|\begin{pmatrix} 1 & 1 & 0 \\ -1 & 6 & 0 \\ 0 & 4 & 0 \end{pmatrix}\right| = 0.$$

An alternative method for the calculation of the determinant of matrices of order higher than three consists of applying elementary operations for determinants to the initial matrix in order to simplify it and be able to obtain the determinant with a smaller number of calculations. In practice, for the simplification of a determinant one can, with the necessary precautions, use the procedure for reduction of matrices seen in Section 4.3.2 with the exception that in the case of determinants we must take into account the following points:

- The multiplication of a row or column by a non-zero number will cause the value of the determinant to change but in its place one can apply part 1 of **Properties 4**.
- The modification of the order of the rows or columns implies a change of sign of the determinant as indicated in part 2 of **Properties 4**.
- Once a row or column has been canceled using a pivot, we will expand the determinant along that row or column to obtain a determinant of smaller size.

Let's see some examples of this method next:

Examples 169.
1)

$$\begin{aligned} &\left|\begin{pmatrix} 1 & -1 & 2 & 3 \\ 3 & -1 & 0 & 2 \\ 1 & 1 & 4 & 6 \\ 0 & -1 & 1 & 0 \end{pmatrix}\right| = \\ &= \text{(C2=C2+C3)} = \left|\begin{pmatrix} 1 & 1 & 2 & 3 \\ 3 & -1 & 0 & 2 \\ 1 & 5 & 4 & 6 \\ 0 & 0 & 1 & 0 \end{pmatrix}\right| \\ &= \left(\begin{matrix}\text{expanding along} \\ \text{row 4}\end{matrix}\right) = (-1)^{4+3} \cdot \left|\begin{pmatrix} 1 & 1 & 3 \\ 3 & -1 & 2 \\ 1 & 5 & 6 \end{pmatrix}\right| \\ &= -\big((-6+45+2)-(-3+10+18)\big) = -16. \end{aligned}$$

2)

$$\left|\begin{pmatrix} -1 & 2 & -1 & 2 & 1 \\ 2 & 3 & 1 & 4 & 1 \\ 2 & 1 & 1 & 4 & 3 \\ 6 & 2 & 1 & -1 & 9 \\ 1 & -2 & 3 & 6 & 4 \end{pmatrix}\right| =$$

$$= \begin{pmatrix} \text{F2=F2+F1} \\ \text{F3=F3+F1} \\ \text{F4=F4+F1} \\ \text{F5=F5+3F1} \end{pmatrix} = \left|\begin{pmatrix} -1 & 2 & -1 & 2 & 1 \\ 1 & 5 & 0 & 6 & 2 \\ 1 & 3 & 0 & 6 & 4 \\ 5 & 4 & 0 & 1 & 10 \\ -2 & 4 & 0 & 12 & 7 \end{pmatrix}\right|$$

$$= \begin{pmatrix} \text{expanding along} \\ \text{column 3} \end{pmatrix} = (-1)\cdot(-1)^{1+3}\cdot \left|\begin{pmatrix} 1 & 5 & 6 & 2 \\ 1 & 3 & 6 & 4 \\ 5 & 4 & 1 & 10 \\ -2 & 4 & 12 & 7 \end{pmatrix}\right|$$

$$= (\text{F2=F2-F1}) = -\left|\begin{pmatrix} 1 & 5 & 6 & 2 \\ 0 & -2 & 0 & 2 \\ 5 & 4 & 1 & 10 \\ -2 & 4 & 12 & 7 \end{pmatrix}\right|$$

$$= (\text{C2=C2+C4}) = -\left|\begin{pmatrix} 1 & 7 & 6 & 2 \\ 0 & 0 & 0 & 2 \\ 5 & 14 & 1 & 10 \\ -2 & 11 & 12 & 7 \end{pmatrix}\right|$$

$$= \begin{pmatrix} \text{expanding along} \\ \text{row 2} \end{pmatrix} = -2\cdot(-1)^{2+4}\cdot \left|\begin{pmatrix} 1 & 7 & 6 \\ 5 & 14 & 1 \\ -2 & 11 & 12 \end{pmatrix}\right|$$

$$= -2\cdot\big((168+330-14)-(-168+11+420)\big) = -2\cdot 221 = -442.$$

4.4.1 Calculation of the inverse via determinants

In this section we will study an alternative method for the calculation of the inverse of a matrix based on the concept of determinant.

Definition 170. Given $A \in \mathcal{M}_n$, we call the adjugate matrix of A and denote it by $Adj(A)$, the matrix:

$$Adj(A) = (\Delta_{ij})_{n\times n} \in \mathcal{M}_n.$$

The adjugate matrix of A is, therefore, the matrix formed by all the cofactors of A arranged in order.

Property 171 (Calculation of the inverse via determinants)**.** *Let* $A \in \mathcal{M}_n$ *such that* $\det(A) \neq 0$*, then* A *is a regular matrix and also:*

$$A^{-1} = \frac{1}{|A|}\cdot(\mathrm{Adj}(A))^t.$$

Examples 172.

1) Let $A = \begin{pmatrix} 1 & -2 \\ 4 & 5 \end{pmatrix}$, then

$$|A| = 5 + 8 = 13 \neq 0,$$

so A is a regular matrix and we can calculate its inverse for which we will first obtain the cofactors of the matrix

$$\begin{array}{ll} \Delta_{11} = (-1)^{1+1}|(5)| = 5, & \Delta_{12} = (-1)^{1+2}|(4)| = -4, \\ \Delta_{21} = (-1)^{2+1}|(-2)| = 2, & \Delta_{22} = (-1)^{2+2}|(1)| = 1 \end{array}$$

and then the adjugate matrix will be:

$$\text{Adj}(A) = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{pmatrix} = \begin{pmatrix} 5 & -4 \\ 2 & 1 \end{pmatrix}.$$

Finally the inverse is:

$$A^{-1} = \frac{1}{|A|} \cdot \text{Adj}(A)^t = \frac{1}{13} \begin{pmatrix} 5 & -4 \\ 2 & 1 \end{pmatrix}^t = \frac{1}{13} \begin{pmatrix} 5 & 2 \\ -4 & 1 \end{pmatrix}.$$

2) Let $M = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ be a generic 2×2 matrix. Then we know that:

- $|M| = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$.
- $\left\{ \begin{array}{ll} \Delta_{11} = a_{22} & \Delta_{12} = -a_{21} \\ \Delta_{21} = -a_{12} & \Delta_{22} = a_{11} \end{array} \right.$.

Therefore the adjugate matrix of M will be

$$\text{Adj}(M) = \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}$$

and then the inverse matrix is

$$M^{-1} = \frac{1}{|M|} \cdot \text{Adj}(M)^t = \frac{1}{a_{11} \cdot a_{22} - a_{12} \cdot a_{21}} \cdot \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}^t$$

so we finally obtain the following formula for the inverse of a 2×2 matrix:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11} \cdot a_{22} - a_{12} \cdot a_{21}} \cdot \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

3) Let $A = \begin{pmatrix} 1 & 6 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ then its determinant is

$$|A| = -2 - 1 - 6 = -9$$

and also its cofactors are

$$\left\{ \begin{array}{lll} \Delta_{11} = \left| \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \right| = -3, & \Delta_{12} = - \left| \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \right| = -1, & \Delta_{13} = \left| \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix} \right| = -1, \\ \Delta_{21} = - \left| \begin{pmatrix} 6 & 0 \\ 1 & -1 \end{pmatrix} \right| = 6, & \Delta_{22} = \left| \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right| = -1, & \Delta_{23} = - \left| \begin{pmatrix} 1 & 6 \\ 0 & 1 \end{pmatrix} \right| = -1, \\ \Delta_{31} = \left| \begin{pmatrix} 6 & 0 \\ 2 & 1 \end{pmatrix} \right| = 6, & \Delta_{32} = - \left| \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \right| = -1, & \Delta_{33} = \left| \begin{pmatrix} 1 & 6 \\ -1 & 2 \end{pmatrix} \right| = 8 \end{array} \right.$$

so its adjugate matrix will be:

$$\text{Adj}(A) = \begin{pmatrix} -3 & -1 & -1 \\ 6 & -1 & -1 \\ 6 & -1 & 8 \end{pmatrix}$$

and A^{-1} will be calculated as:

$$A^{-1} = \frac{1}{|A|} \cdot \text{Adj}(A)^t = \frac{1}{-9} \cdot \begin{pmatrix} -3 & 6 & 6 \\ -1 & -1 & -1 \\ -1 & -1 & 8 \end{pmatrix}.$$

Aquí está la traducción al inglés del fragmento final del capítulo de matrices, cumpliendo estrictamente con todas las indicaciones.

"'latex

4.5 Additional Material

4.5.1 Rows and columns of proportions

Extension of concepts about product of a number by a matrix. Page 126

We have seen that when several data a_1, a_2,..., a_n intervene in a certain phenomenon, we can represent this information by an element of $\mathbb{R}^n$ in the form,

$$(a_1, a_2, \ldots, a_n).$$

It is frequent that it is of interest to determine what percentage each quantity represents with respect to the total. This can be done using percentages or fractions (rates per one):

- The calculation of percentages for each quantity of $(a_1, a_2, \ldots, a_n)$ is performed in the following way:
  - Percentage of a_1 with respect to the total= $\dfrac{100}{a_1 + a_2 + \cdots + a_n} a_1\%$.
  - Percentage of a_2 with respect to the total= $\dfrac{100}{a_1 + a_2 + \cdots + a_n} a_2\%$.
  - $\vdots$
  - Percentage of a_n with respect to the total= $\dfrac{100}{a_1 + a_2 + \cdots + a_n} a_n\%$.
- The calculation of fractions (rates per one) for each quantity of $(a_1, a_2, \ldots, a_n)$ is performed as follows:
  - Fraction (rate per one) of a_1 with respect to the total= $\dfrac{1}{a_1 + a_2 + \cdots + a_n} a_1$.
  - Fraction (rate per one) of a_2 with respect to the total= $\dfrac{1}{a_1 + a_2 + \cdots + a_n} a_2$.
  - $\vdots$
  - Fraction (rate per one) of a_n with respect to the total= $\dfrac{1}{a_1 + a_2 + \cdots + a_n} a_n$.

If we represent the percentages corresponding to $(a_1, a_2, \ldots, a_n)$ by an element of $\mathbb{R}^n$ and use the definition of the product of a number by a matrix we have,

$$\left(\frac{100}{a_1 + a_2 + \cdots + a_n} a_1, \frac{100}{a_1 + a_2 + \cdots + a_n} a_2, \ldots, \frac{100}{a_1 + a_2 + \cdots + a_n} a_n\right) =$$

$$= \frac{100}{a_1 + a_2 + \cdots + a_n}(a_1, a_2, \ldots, a_n).$$

For its part, if we do the same for the fractions (rates per one) we obtain

$$(\frac{1}{a_1 + a_2 + \cdots + a_n}a_1, \frac{1}{a_1 + a_2 + \cdots + a_n}a_2, \ldots, \frac{1}{a_1 + a_2 + \cdots + a_n}a_n) =$$

$$= \frac{1}{a_1 + a_2 + \cdots + a_n}(a_1, a_2, \ldots, a_n).$$

In other words, given the distribution of quantities $(a_1, a_2, \ldots, a_n)$, we have that

The tuple of corresponding percentages is:	$\frac{100}{a_1 + a_2 + \cdots + a_n}(a_1, a_2, \ldots, a_n)$
The tuple of corresponding fractions (rates per one) is:	$\frac{1}{a_1 + a_2 + \cdots + a_n}(a_1, a_2, \ldots, a_n)$

Ejemplo 5. The consumption of raw materials in a certain industrial zone is given by: 1123 tons of steel, 820 tons of aluminum, 530 tons of plastic materials. We can represent this distribution by the 3-tuple

$$(1123, 820, 530) \in \mathbb{R}^3.$$

Let's calculate the percentages that each species represents with respect to the total:

- Percentage of steel $= 100\frac{1123}{1123 + 820 + 530} = 45.41\%$.
- Percentage of aluminum $= 100\frac{820}{1123 + 820 + 530} = 33.15\%$.
- Percentage of plastic materials $= 100\frac{530}{1123 + 820 + 530} = 21.43\%$.

As we have just seen, using the definition of the product of a number by a matrix we can obtain these same percentages through the operation

$$\frac{100}{1123 + 820 + 530}(1123, 820, 530) = (45.41, 33.15, 21.43).$$

In the same way the tuple of fractions (rates per one) will be,

$$\frac{1}{1123 + 820 + 530}(1123, 820, 530) = (0.4541, 0.3315, 0.2143).$$

If we had a different distribution of raw material consumptions the percentages could vary. For example if the distribution is represented by $(120, 201, 150)$ the percentages will be,

$$\frac{100}{120 + 201 + 150}(120, 201, 150) = (25.47, 42.67, 31.84).$$

We see that the distributions $(1123, 820, 530)$ and $(120, 201, 150)$ lead to different percentages. However, we can find different distributions that give rise to the same percentages. For example, the percentages corresponding to the distribution $(2246, 1640, 1060)$ are

$$\frac{100}{2246 + 1640 + 1060}(2246, 1640, 1060) = (45.41, 33.15, 21.43).$$

4.6 Simplification rules and matrix equalities

Extension on operations and simplifications with matrices. Page 137

In different points of this chapter and the following ones we will encounter expressions and equalities that involve operations with matrices. We know that, when handling expressions or equalities with numbers, there exist rules that allow their simplification. In the following property we summarize the rules of simplification that we can apply in the matrix case. It can be seen that many of them are equivalent to those we use for numerical expressions but others require special care.

Propiedades 6.

1. *Let $A, B, C \in \mathcal{M}_{m\times n}$, then:*

$$\begin{aligned} A+B=A+C \quad &\Rightarrow \quad B=C, \\ A+B=C \quad &\Rightarrow \quad B=C-A. \end{aligned}$$

2. *Let $A \in \mathcal{M}_n$ be a regular matrix and $B, C \in \mathcal{M}_{m\times n}$, then:*

$$\begin{aligned} B\cdot A=C\cdot A \quad &\Rightarrow \quad B=C, \\ B\cdot A=C \quad &\Rightarrow \quad B=C\cdot A^{-1}. \end{aligned}$$

3. *Let $A \in \mathcal{M}_m$ be a regular matrix and $B, C \in \mathcal{M}_{m\times n}$, then:*

$$\begin{aligned} A\cdot B=A\cdot C \quad &\Rightarrow \quad B=C, \\ A\cdot B=C \quad &\Rightarrow \quad B=A^{-1}\cdot C. \end{aligned}$$

4. *Let $r \in \mathbb{R}$, $r \neq 0$ and $B, C \in \mathcal{M}_{m\times n}$, then:*

$$\begin{aligned} r\cdot B=r\cdot C \quad &\Rightarrow \quad B=C, \\ r\cdot B=C \quad &\Rightarrow \quad B=\frac{1}{r}\cdot C. \end{aligned}$$

Proof. The proof of all points is simple using all the properties we already know about matrices. □

Remark.
• Note that in properties 2 and 3 it is fundamental that the condition that the matrix A is regular is fulfilled. If A is a singular matrix we cannot guarantee that it is possible to simplify it nor that it is possible to solve for it. For example, it is easily verified that

$$\underbrace{\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}\cdot\begin{pmatrix}1 & 2\\ -1 & -2\end{pmatrix}}_{=0_{2\times 2}} = \underbrace{\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}\cdot\begin{pmatrix}-3 & 1\\ 3 & -1\end{pmatrix}}_{=0_{2\times 2}}$$

however, we cannot simplify in this equality the matrix $\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$ because if we do so we obtain

$$\cancel{\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}}\cdot\begin{pmatrix}1 & 2\\ -1 & -2\end{pmatrix} = \cancel{\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}}\cdot\begin{pmatrix}-3 & 1\\ 3 & -1\end{pmatrix} \Rightarrow \begin{pmatrix}1 & 2\\ -1 & -2\end{pmatrix} = \begin{pmatrix}-3 & 1\\ 3 & -1\end{pmatrix}$$

which evidently is not true. This is due to the fact that, as we know, the simplified matrix is not regular.

• There are other situations in which the usual rules for the manipulation of numbers cannot be applied when we work with matrices. Thus for example, if we take the matrices $A, B \in \mathcal{M}_n$ and we try to calculate the product $A \cdot B \cdot A^{-1}$, we cannot perform the simplification

$$\not{A} \cdot B \cdot \not{A}^{-1} = B$$

because we can eliminate A with its inverse, A^{-1}, only when they are next to each other, which is not the case here. Note also that, since the product of matrices does not have the commutative property, in $A \cdot B \cdot A^{-1}$ we cannot alter the order to get $B \cdot A \cdot A^{-1}$ and thus be able to simplify.

Ejemplos 7.
1) *Given two matrices $A, B \in \mathcal{M}_n$, check if the equality holds*

$$(A + B)^2 = A^2 + B^2 + 2 \cdot A \cdot B.$$

We have that

$$\begin{aligned} (A + B)^2 &= (A + B) \cdot (A + B) = A \cdot (A + B) + B \cdot (A + B) \\ &= A \cdot A + A \cdot B + B \cdot A + B \cdot B = A^2 + B^2 + A \cdot B + B \cdot A \end{aligned}$$

and then

$$\begin{aligned} (A + B)^2 &= A^2 + B^2 + 2 \cdot A \cdot B \\ &\Updownarrow \\ A^2 + B^2 + A \cdot B + B \cdot A &= A^2 + B^2 + 2 \cdot A \cdot B \\ &\Updownarrow \\ A \cdot B + B \cdot A &= 2 \cdot A \cdot B \\ &\Updownarrow \\ A \cdot B + B \cdot A &= A \cdot B + A \cdot B \\ &\Updownarrow \\ B \cdot A &= A \cdot B. \end{aligned}$$

Therefore the equality proposed in the example will be correct only if it holds that $B \cdot A = A \cdot B$ which in general is not true. The equality will hold only for those matrices A and B that make the equality true

$$B \cdot A = A \cdot B.$$

2) *Simplify the expression $(A + B)^2 - A \cdot (A + B) - (B + A) \cdot B$.*

Using some of the equalities obtained in the previous exercise we have:

$$\begin{aligned} &(A + B)^2 - A \cdot (A + B) - (B + A) \cdot B \\ &= A^2 + B^2 + A \cdot B + B \cdot A - A^2 - A \cdot B - B^2 - A \cdot B \\ &= B \cdot A - A \cdot B. \end{aligned}$$

3) *Simplify the expression*

$$-2 \cdot A^t \cdot B - (B^t \cdot A)^t + \left[B^t \cdot A + (A^t \cdot B^t)^t\right]^t.$$

We have that

$$\begin{aligned}
&-2\cdot A^t\cdot B-(B^t\cdot A)^t+\left[B^t\cdot A+(A^t\cdot B^t)^t\right]^t\\
&\quad= -2\cdot A^t\cdot B-A^t\cdot B^{tt}+(B^t\cdot A)^t+((A^t\cdot B^t)^t)^t\\
&\quad= -2\cdot A^t\cdot B-A^t\cdot B+A^t\cdot B^{tt}+A^t\cdot B^t\\
&\quad= A^t\cdot B^t-2\cdot A^t\cdot B=A^t\cdot(B^t-2\cdot B).
\end{aligned}$$

4) *Calculate the matrices $X, Y \in \mathcal{M}_2$ that satisfy the following equalities:*

$$\begin{cases} 2\cdot X-5\cdot Y = \begin{pmatrix}1&-2\\0&1\end{pmatrix}\\ -X+3\cdot Y=\begin{pmatrix}2&1\\3&3\end{pmatrix}\end{cases}.$$

If we multiply the second equation by 2 we obtain

$$\begin{cases} 2\cdot X-5\cdot Y = \begin{pmatrix}1&-2\\0&1\end{pmatrix}\\ -2\cdot X+6\cdot Y=2\cdot\begin{pmatrix}2&1\\3&3\end{pmatrix}\end{cases},$$

adding both equations

$$Y=\begin{pmatrix}1&-2\\0&1\end{pmatrix}+2\cdot\begin{pmatrix}2&1\\3&3\end{pmatrix}=\begin{pmatrix}5&0\\6&7\end{pmatrix}$$

and finally, using the second of the equations we had at the beginning:

$$X=3\cdot Y-\begin{pmatrix}2&1\\3&3\end{pmatrix}=\begin{pmatrix}15&0\\18&21\end{pmatrix}-\begin{pmatrix}2&1\\3&3\end{pmatrix}=\begin{pmatrix}13&-1\\15&18\end{pmatrix}.$$

5) *Calculate the matrices C_1 and C_2 that satisfy:*

$$\textbf{a)}\ \begin{pmatrix}1&1\\1&0\end{pmatrix}\cdot C_1=\begin{pmatrix}2&0\\1&1\end{pmatrix},\qquad \textbf{b)}\ C_2\cdot\begin{pmatrix}1&1\\1&0\end{pmatrix}=\begin{pmatrix}2&0\\1&1\end{pmatrix}.$$

Using exercise 1 from page 134 we know that the matrix $\begin{pmatrix}1&1\\1&0\end{pmatrix}$ is regular and that its inverse is:

$$\begin{pmatrix}1&1\\1&0\end{pmatrix}^{-1}=\begin{pmatrix}0&1\\1&-1\end{pmatrix}.$$

Since the matrix is regular we can solve for it obtaining:

$$C_1=\begin{pmatrix}1&1\\1&0\end{pmatrix}^{-1}\cdot\begin{pmatrix}2&0\\1&1\end{pmatrix}=\begin{pmatrix}0&1\\1&-1\end{pmatrix}\cdot\begin{pmatrix}2&0\\1&1\end{pmatrix}=\begin{pmatrix}1&1\\1&-1\end{pmatrix},$$

$$C_2=\begin{pmatrix}2&0\\1&1\end{pmatrix}\cdot\begin{pmatrix}1&1\\1&0\end{pmatrix}^{-1}=\begin{pmatrix}2&0\\1&1\end{pmatrix}\cdot\begin{pmatrix}0&1\\1&-1\end{pmatrix}=\begin{pmatrix}0&2\\1&0\end{pmatrix}.$$

4.7 Matrix models based on powers of matrices

Extension of concepts about powers of matrices. Page 137

The product and the power of matrices are fundamental in the formulation of the most important matrix models.

Suppose we are studying a phenomenon in which several magnitudes $a_1, a_2, \ldots, a_k$ intervene that vary with respect to time. If we study that phenomenon over several periods, $n = 0$ (initial period), $n = 1$, $n = 2$, etc., the magnitudes $a_1, a_2, \ldots, a_k$ will take different values. If we arrange in tuple form the value of the magnitudes in each period n, we will obtain a list of k-tuples, P_0, P_1, $\ldots$, P_n that provide us with the information of the phenomenon in each period.

The important thing here would be to be able to calculate the k-tuples corresponding to future periods so that we can predict the evolution of the phenomenon. It is then when the calculation of matrix powers comes into play because in numerous situations, if we know the initial situation of the phenomenon, that is, we know the tuple P_0 corresponding to the initial period $n = 0$, we can calculate the tuple of any period n by means of a formula of the type

$$P_n = A^n \cdot P_0,$$

where A is a square matrix of order k which is called the **transition matrix** and which governs the changes that the phenomenon experiences from one period to the next.

We will illustrate this better next with a classic example of a matrix model based on the powering of matrices.

Ejemplo 8. Suppose that in a certain commercial sector three companies compete, which we will call A, B and C. From one year to the next, the customers of each of them decide to remain loyal or switch to one of the others. A study is carried out on the movements between the three companies and it is observed that year after year the customers show a similar behavior determined by the data in the following table:

	Customers of A	**Customers of B**	**Customers of C**
Switch to become customers of A	80%	10%	10%
Switch to become customers of B	10%	60%	20%
Switch to become customers of C	10%	30%	70%

For example, we see that each year 80% of the customers of A remain loyal to A, 10% switch to B and 10% to C.

Suppose also that in the year the studies began, the company A had 210 customers, B had 190 and C, 320.

We are going to set up a matrix model to study this problem. Assuming that the year $k = 0$ is the year the study of the customers of the three companies began, we will call:

- A_k = number of customers of company A after k years.
- B_k = number of customers of company B after k years.
- C_k = number of customers of company C after k years.

The information for each year will be grouped into a column tuple which we will denote as P_k,

$$P_k = \begin{pmatrix} A_k \\ B_k \\ C_k \end{pmatrix}.$$

According to the problem data we have that $A_0 = 210$, $B_0 = 190$ and $C_0 = 320$ so that

$$P_0 = \begin{pmatrix} 210 \\ 190 \\ 320 \end{pmatrix}.$$

Applying the transition table it is easy to calculate the customers that will be in each company if we know those that were there the previous year. Thus, if in year k we have A_k in A, B_k in B and C_k in C, in year $k+1$ we will have:

- $\underbrace{A_{k+1}}_{\text{customers in A in year } k+1} = 80\% \text{ of } A_k + 10\% \text{ of } B_k + 10\% \text{ of } C_k = 0.8A_k + 0.1B_k + 0.1C_k.$

- $\underbrace{B_{k+1}}_{\text{customers in B in year } k+1} = 10\% \text{ of } A_k + 60\% \text{ of } B_k + 20\% \text{ of } C_k = 0.1A_k + 0.6B_k + 0.2C_k.$

- $\underbrace{C_{k+1}}_{\text{customers in C in year } k+1} = 10\% \text{ of } A_k + 30\% \text{ of } B_k + 70\% \text{ of } C_k = 0.1A_k + 0.3B_k + 0.7C_k$

Writing all this information in a column and using the definition of the product of matrices, it is easy to realize that

$$P_{k+1} = \begin{pmatrix} 0.8A_k + 0.1B_k + 0.1C_k \\ 0.1A_k + 0.6B_k + 0.2C_k \\ 0.1A_k + 0.3B_k + 0.7C_k \end{pmatrix} = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} A_k \\ B_k \\ C_k \end{pmatrix} = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \cdot P_k.$$

Calling $A = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix}$, in short we have proven that

$$P_{k+1} = AP_k.$$

We therefore have,

$$\begin{array}{rcl} P_1 & = & AP_0 \\ P_2 & = & AP_1 \\ P_3 & = & AP_2 \\ P_4 & = & AP_3 \\ & \text{etc.} & \end{array}$$

Then, if we want to calculate P_4 according to this scheme, since the only data we know are those of the initial year, i.e. P_0, we will have to calculate first P_1, then P_2 and P_3 and finally P_4. Now, we have that

$$\begin{array}{rcl} P_2 & = & AP_1 = A(AP_0) = (AA)P_0 = A^2P_0. \\ P_3 & = & AP_2 = (\text{using the previous equation}) = A(A^2P_0) = (AA^2)P_0 = A^3P_0. \\ P_4 & = & AP_3 = (\text{using the previous equation}) = A(A^3P_0) = (AA^3)P_0 = A^4P_0. \end{array}$$

Then using powers of matrices we can calculate P_4 without needing to obtain beforehand P_0, P_1, P_2 and P_3. In reality, it is evident that this process can be applied iteratively as many times as we want so that, in general,

$$\boxed{P_k = A^kP_0}\,. \tag{4.5}$$

What we see here is that the distribution of customers in year k, P_k, is determined by the initial distribution, P_0 and the k-th power of A. The matrix A regulates the passage from one year to the next and is the **transition matrix** for this problem.

Since we know the initial distribution of customers, P_0, we can easily calculate the distribution in successive years. To do this, we calculate several powers of A:

$$\begin{aligned}
A^2 = AA &= \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} = \begin{pmatrix} 0.66 & 0.17 & 0.17 \\ 0.16 & 0.43 & 0.27 \\ 0.18 & 0.4 & 0.56 \end{pmatrix}. \\
A^3 = AA^2 &= \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 0.66 & 0.17 & 0.17 \\ 0.16 & 0.43 & 0.27 \\ 0.18 & 0.4 & 0.56 \end{pmatrix} = \begin{pmatrix} 0.562 & 0.219 & 0.219 \\ 0.198 & 0.355 & 0.291 \\ 0.24 & 0.426 & 0.49 \end{pmatrix}. \\
A^4 = AA^3 &= \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 0.562 & 0.219 & 0.219 \\ 0.198 & 0.355 & 0.291 \\ 0.24 & 0.426 & 0.49 \end{pmatrix} = \begin{pmatrix} 0.4934 & 0.2533 & 0.2533 \\ 0.223 & 0.3201 & 0.2945 \\ 0.2836 & 0.4266 & 0.4522 \end{pmatrix}.
\end{aligned}$$

Using these calculations with equation (4.5) we have that

$$\begin{aligned}
P_1 = AP_0 &= \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 210 \\ 190 \\ 320 \end{pmatrix} = \begin{pmatrix} 219 \\ 199 \\ 302 \end{pmatrix}. \\
P_2 = A^2 P_0 &= \begin{pmatrix} 0.66 & 0.17 & 0.17 \\ 0.16 & 0.43 & 0.27 \\ 0.18 & 0.4 & 0.56 \end{pmatrix} \begin{pmatrix} 210 \\ 190 \\ 320 \end{pmatrix} = \begin{pmatrix} 225.3 \\ 201.7 \\ 293 \end{pmatrix}. \\
P_3 = A^3 P_0 &= \begin{pmatrix} 0.562 & 0.219 & 0.219 \\ 0.198 & 0.355 & 0.291 \\ 0.24 & 0.426 & 0.49 \end{pmatrix} \begin{pmatrix} 210 \\ 190 \\ 320 \end{pmatrix} = \begin{pmatrix} 229.71 \\ 202.15 \\ 288.14 \end{pmatrix}. \\
P_4 = A^4 P_0 &= \begin{pmatrix} 0.4934 & 0.2533 & 0.2533 \\ 0.223 & 0.3201 & 0.2945 \\ 0.2836 & 0.4266 & 0.4522 \end{pmatrix} \begin{pmatrix} 210 \\ 190 \\ 320 \end{pmatrix} = \begin{pmatrix} 232.797 \\ 201.889 \\ 285.314 \end{pmatrix}.
\end{aligned}$$

After performing all these operations, it is evident that the greatest obstacle is the calculation of the powers of A due to the fact that it is not a diagonal matrix. Up to the fourth power the calculation could be done manually but if we wanted higher powers like A^{20} or A^{30}, it seems essential to resort to other techniques. On the other hand, once this model is set up, several issues arise to be solved:

a) Is it possible to study the future trend in the distribution of customers? Even more interesting than calculating the customers in a specific year would be to be able to describe the future behavior by determining if the customers tend to choose with greater priority one of the three companies or if on the contrary they distribute themselves in a homogeneous way among them.

b) Do equilibrium distributions exist? For example, the presidents of the three companies could try to agree to distribute the customer market so that it remains constant from one year to the next. For this, we should choose an initial distribution of customers

$$P_0 = \begin{pmatrix} A_0 \\ B_0 \\ C_0 \end{pmatrix}$$

such that the distributions in subsequent years, P_1, P_2, P_3, etc. are always equal. If the distributions of customers in year zero and in year one are equal we will have $P_0 = P_1$ and since we know that $P_1 = AP_0$ we deduce that

$$\boxed{AP_0 = P_0}\,.$$

If we find an initial distribution, P_0, that satisfies this last condition, it is not difficult to check that the distribution in all subsequent years is always the same since

$$P_1 = AP_0 = P_0,$$

$$\begin{aligned} P_2 &= AP_1 = AP_0 = P_0, \\ P_3 &= AP_2 = AP_0 = P_0, \\ P_4 &= AP_3 = AP_0 = P_0, \\ &\text{etc.} \end{aligned}$$

For example if $A_0 = 600$, $B_0 = 500$, $C_0 = 700$ then $P_0 = (600, 500, 700)$ and it is easy to check that

$$AP_0 = A \begin{pmatrix} 600 \\ 500 \\ 700 \end{pmatrix} = \begin{pmatrix} 600 \\ 700 \\ 500 \end{pmatrix} = P_0.$$

Therefore, since the condition $AP_0 = P_0$ is satisfied, in all subsequent years we will always have the same distribution of customers given by the 3-tuple $(600, 500, 700)$.

c) If the total number of customers is increasing or decreasing each time, it will be impossible for the number of customers of the three companies in subsequent years to remain constant. In such a case, the presidents could agree that at least the percentages of customers for each company are the same in all years. It is simple to prove that the distributions of customers in the initial year, P_0, and in year one, P_1, represent the same percentages if we can find $\lambda \neq 0$ such that $P_1 = \lambda P_0$. Since $P_1 = AP_0$ we conclude that

$$\boxed{AP_0 = \lambda P_0}\ .$$

In this case we ask about the way to calculate the value λ and the initial distribution P_0.
