# COURSE 6

## The rank of a matrix

Let $m, n \in \mathbb{N}^*$, $A = (a_{ij}) \in M_{m,n}(K)$.

**Definition 1.** Let $i_1, \ldots, i_k, j_1, \ldots, j_l \in \mathbb{N}^*$ cu $1 \leq i_1 < \cdots < i_k \leq m$ and $1 \leq j_1 < \cdots < j_l \leq n$. A matrix

$$\begin{pmatrix} a_{i_1j_1} & a_{i_1j_2} & \ldots & a_{i_1j_k} \\ a_{i_2j_1} & a_{i_2j_2} & \ldots & a_{i_2j_k} \\ \vdots & \vdots & & \vdots \\ a_{i_lj_1} & a_{i_lj_2} & \ldots & a_{i_lj_k} \end{pmatrix}$$

formed by taking the elementels of $A$ which are situated at the intersections of the rows $i_1, \ldots, i_k$ with the columns $j_1, \ldots, j_l$ is called $k \times l$ **submatrix of** $A$. The determinant of a $k \times k$ submatrix is called **minor of** $A$ **of order** $k$.

**Definition 2.** Let $A \in M_{m,n}(K)$. If $A$ is not the zero matrix, i.e. $A \neq O_{m,n}$, we say that **the rank of the matrix** $A$ is $r$, and we write $\operatorname{rank} A = r$, if $A$ has a non-zero minor of order $r$ all the minors of $A$ of order greater than $r$ (if they exist) are 0. By definition, $\operatorname{rank} O_{m,n} = 0$.

**Remark 3.** a) $\operatorname{rank} A \leq \min\{m, n\}$.
b) If $A \in M_n(K)$ then $\operatorname{rank} A = n$ dif and only if $\det A \neq 0$.
c) $\operatorname{rank} A = \operatorname{rank} {}^tA$.

For the following part of this section, we take $m, n \in \mathbb{N}^*$, $A = (a_{ij}) \in M_{m,n}(K)$ and $A \neq O_{m,n}$.

Finding the rank of $A$ by definition involves, most of the time, a large number of computations (of minors). The next theorem is a first step for reducing the number of these computations.

**Theorem 4.** $\operatorname{rank} A = r$ if and only if $A$ has a non-zero minor of order $r$ and all $r+1$-size minors of $A$ (if they exist) are 0.

*Proof.* □

**Theorem 5.** The rank of the matrix $A$ is the maximum number of columns (rows) we can choose from the columns (rows) of $A$ such that none of them is a linear combination of the others.

*Proof.* Suppose that the rank of $A$ is $r$. Then $A$ has a non-zero minor of order $r$. For simpler notations, we consider that

$$d = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \ldots & a_{rr} \end{vmatrix} \neq 0$$

and any $r+1$-size minor is zero. (The proof of the general case works in the same way, only the notations are more complicated.) Therefore the determinant

$$D_{ij} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1j} \\ a_{21} & a_{22} & \dots & a_{2r} & a_{2j} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rr} & a_{rj} \\ a_{i1} & a_{i2} & \dots & a_{ir} & a_{ij} \end{vmatrix}$$

of size $r+1$ resulted by adding to $d$ the $i$'th row and $j$'th column of $A$ ($1 \le i \le m$, $r < j \le n$) is zero, i.e. $D_{ij} = 0$. Notice that if $1 \le i \le r$ then $D_{ij}$ has two equal rows, and if $r < i \le m$ and $r < j \le n$ then $D_{ij}$ is a $r+1$-size minor of $A$ resulted by adding to $d$ the row $i$ and the column $j$. Expanding $D_{ij}$ along the row $r+1$, we get

$$a_{i1}d_1 + a_{i2}d_2 + \cdots + a_{ir}d_r + a_{ij}d = 0$$

where the cofactors $d_1, d_2, \dots, d_r$ do not depend on the added row $i$. It follows that

$$a_{ij} = -d^{-1}d_1a_{i1} - d^{-1}d_2a_{i2} - \cdots - d^{-1}d_ra_{ir}$$

for all $i = 1, 2, \dots, m$ and $j = r+1, \dots, n$ thus

$$c_j = \alpha_1c_1 + \alpha_2c_2 + \cdots + \alpha_rc_r \text{ for all } j = r+1, \dots, n,$$

where $\alpha_k = -d^{-1}d_k$, $1 \le k \le r$, i.e. $c_j$ is a linear combination of $c_1, c_2, \dots, c_r$.

This way we proved that the maximum number of columns we can choose from the columns of $A$ such that none of them is a linear combination of the others is at most $r$. If this number is strictly smaller than $r$, then one of $c_1, \dots, c_r$ will be a linear combination of the others and $d = 0$, which is not possible.

Thus the maximum number of columns we can choose from the columns of $A$ such that none of them is a linear combination of the others is exactly $r$ and the proof is now complete. $\square$

**Corollary 6.** $\operatorname{rank} A = r$ if and only if $A$ has a non-zero minor $d$ of order $r$ and all the other rows (columns) of $A$ are linear combinations of the the rows (columns) of $A$ whose elements are the entries of $d$.

**Corollary 7.** If $m, n, p \in \mathbb{N}^*$, $A = (a_{ij}) \in M_{m,n}(K)$ and $B = (b_{ij}) \in M_{n,p}(K)$ ($K$ field) then $\operatorname{rank}(AB) \le \min\{\operatorname{rank} A,\ \operatorname{rank} B\}$.

If one of the given matrices is zero, the property is obvious. So, let us consider both our matrices non-zero and let us suppose that $\min\{\operatorname{rank} A,\ \operatorname{rank} B\} = \operatorname{rank} B = r \in \mathbb{N}^*$ and that a non-zero minor of $B$ of size $r$ can be extracted from the columns $j_1, \dots, j_r$ with $1 \le j_1 < \cdots < j_r \le p$. (For the other case, one can rephrase the statement for the transposes of our matrices, then one can use the same reasoning to find the expected result.) The columns of $AB$ are

$$A\begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix},\ A\begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{pmatrix}, \dots,\ A\begin{pmatrix} b_{1p} \\ b_{2p} \\ \vdots \\ b_{np} \end{pmatrix}.$$

From corollary 6 we deduce that for any $k \in \{1, \ldots, p\} \setminus \{j_1, \ldots, j_r\}$, there exist $\alpha_{1k}, \ldots, \alpha_{rk} \in K$ such that

$$\begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{pmatrix} = \alpha_{1k} \begin{pmatrix} b_{1j_1} \\ b_{2j_1} \\ \vdots \\ b_{nj_1} \end{pmatrix} + \alpha_{2k} \begin{pmatrix} b_{1j_2} \\ b_{2j_2} \\ \vdots \\ b_{nj_2} \end{pmatrix} + \cdots + \alpha_{rk} \begin{pmatrix} b_{1j_r} \\ b_{2j_r} \\ \vdots \\ b_{nj_r} \end{pmatrix}.$$

Hence

$$A \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{pmatrix} = \alpha_{1k} \cdot A \begin{pmatrix} b_{1j_1} \\ b_{2j_1} \\ \vdots \\ b_{nj_1} \end{pmatrix} + \alpha_{2k} A \cdot \begin{pmatrix} b_{1j_2} \\ b_{2j_2} \\ \vdots \\ b_{nj_2} \end{pmatrix} + \cdots + \alpha_{rk} \cdot A \begin{pmatrix} b_{1j_r} \\ b_{2j_r} \\ \vdots \\ b_{nj_r} \end{pmatrix},$$

which means that in $AB$ all the columns $k \in \{1, \ldots, p\} \setminus \{j_1, \ldots, j_r\}$ are linear combinations of the columns $j_1, \ldots, j_r$. Thus the rank of the matrix $AB$ is at most $r$.

**Corollary 8.** Let $n \in \mathbb{N}^*$ and $K$ be a field. A matrix $A \in M_n(K)$ is invertible (i.e. a unit in $(M_n(K), +, \cdot)$) if and only if $\det A \neq 0$.

**Corollary 9.** $\operatorname{rank} A = r$ if and only if there exists a non-zero minor $d$ of $A$ of order $r$ and all the $r+1$-size minors of $A$ resulted by adding one of remained rows and columns to $d$ are 0 (if they exist, of course).

**An algorithm for finding the rank of a matrix:**
Corollary 9 shows that for a matrix $A \neq O_{m,n}$, $\operatorname{rank} A$ can be determined in the following way: we start with a non-zero minor $d$ of $A$ and we compute all the minors of $A$ obtained by adding $d$ one of the remained rows and one of the remained columns until we find a non-zero minor, minor which will be the subject of a similar approach. In finitely many steps, we will find a non-zero minor of order $r$ of $A$ for which all the $r+1$-size minors resulted by adding it one of remained rows and columns are zero. Thus $r = \operatorname{rank} A$.

# Systems of linear equations

Let $K$ be a field and let us consider the system of $m$ linear equations with $n$ unknowns:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1}$$

where $a_{ij}, b_j \in K$, $i = 1, \ldots, m$; $j = 1, \ldots, n$. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \ \overline{A} = \begin{pmatrix} a_{11} & \ldots & a_{1n} & b_1 \\ a_{21} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mn} & b_m \end{pmatrix} \text{ and } B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

We remind that $A \in M_{m,n}(K)$ is **the matrix of the system** (1), $B$ is **the matrix of constant terms** and $\overline{A}$ is **the augmented matrix of the system**. If all the constant terms are zero, i.e. $b_1 = b_2 = \cdots = b_m = 0$, the system (1) is a **homogeneous linear system**. By denoting

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

the system (1) can be written as a matrix equation

$$AX = B \tag{2}$$

The system $AX = O_{m,1}$ is **the homogeneous system associated** to the system $AX = B$.

**Definition 10.** An $n$-tuple $(\alpha_1, \ldots, \alpha_n) \in K^n$ is a **solution of the system** (1) if the all the equalities resulted by replacing $x_i$ with $\alpha_i$ $(i = 1, \ldots, n)$ in (1) are true. The system (1) is called **consistent** if it has at least one solution. Otherwise, the system (1) is **inconsistent**. Two **systems** of linear equations with $n$ unknowns are **equivalent** if they have the same solution set.

**Remarks 11.** a) Cramer's Theorem states that for $m = n$ and $\det A \neq 0$ the system (1) is consistent, with a unique solution, and its solution is given by Cramer's formulas.
b) If (1) is a homogeneous system, then $(0, 0, \ldots, 0) \in K^n$ is a solution of (1), called **the trivial solution**, so any homogeneous linear system is consistent.

The following result is a very important tool for the study of the consistency of the general linear systems.

**Theorem 12. (Kronecker-Capelli)** The linear system (1) is consistent if and only if the rank of its matrix is the same as the rank of its augmented matrix, i.e. $\operatorname{rank} A = \operatorname{rank} \overline{A}$.

*Proof.* □

Let us consider that $\operatorname{rank} A = r$. Based on how one can determine the rank of a matrix one can restate the previous theorem as follows:

**Theorem 13. (Rouché)** Let $d_p$ be a nonzero $r \times r$ minor of the matrix $A$. The system (1) is consistent if and only if all the $(r+1) \times (r+1)$ minors of $\overline{A}$ obtained by completing $d_p$ with a column of constant terms and the corresponding row are zero (if such $(r+1) \times (r+1)$ minors exist).

We end this section by presenting an algorithm for solving arbitrary systems of linear equations based on Rouché Theorem.

**An algorithm for solving systems of linear equations:**
We begin by studying the consistency of (1) by using Rouché's theorem and let us consider that we have found a minor $d_p$ of $A$. If one finds a nonzero $(r+1) \times (r+1)$ minor which completes $d_p$ as in Rouché Theorem, then (1) is inconsistent and the algorithm ends. If $r = m$ or all the Rouché Theorem $(r+1) \times (r+1)$ minor completions of $d_p$ are 0, then (1) is consistent.

We call the unknowns corresponding to the the entries of $d_p$ **main unknowns** and the other unknowns **side unknowns**. For simpler notations, we consider that the minor $d_p$ was "cut" from the first $r$ rows and the first $r$ columns of $A$. One considers only the $r$ equations which determined the rows of $d_p$. Since $\operatorname{rank} \overline{A} = \operatorname{rank} A = r$, Corollary 6 tells us that all the other equations are linear combinations" of these $r$ equations, hence (1) is equivalent to

$$\begin{cases} a_{11}x_1 + x_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + x_{22}x_2 + \cdots + a_{2n}x_n = b_1 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + x_{r2}x_2 + \cdots + a_{rn}x_n = b_r \end{cases} \tag{3}$$

If $n = r$, i.e. all the unknowns are main unknowns, then (3) is a Cramer system. The Cramer's rule gives us its unique solution, hence the unique solution of (1).

Otherwise, $n > r$, and $x_{r+1}, \ldots, x_n$ are side unknowns. We can assign them arbitrary "values" from $K$, $\alpha_{r+1}, \ldots, \alpha_n$, respectively. Then (3) becomes

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1r}x_r = b_1 - a_{1,r+1}\alpha_{r+1} - \cdots - a_{1n}\alpha_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2r}x_r = b_2 - a_{2,r+1}\alpha_{r+1} - \cdots - a_{2n}\alpha_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + a_{r2}x_2 + \cdots + a_{rr}x_r = b_r - a_{r,r+1}\alpha_{r+1} - \cdots - a_{rn}\alpha_n \end{cases} \tag{4}$$

The determinant of the matrix of (4) is $d_p \neq 0$, hence we can express the main unknowns using the side unknowns, by solving the Cramer system (4).

# Elementary operations on matrices. Applications

Let $K$ be a field, $m, n \in \mathbb{N}^*$ and $A = (a_{ij}) \in M_{m,n}(K)$.

**Definition 14.** By an **elementary operation on the rows (columns)** of a matrix we understand one of the following:
(I) the interchange of two rows (columns).
(II) multiplying a row (column) by a non-zero element $\alpha \in K$.
(III) multiplying a row (column) by an element $\alpha \in K$ (also called scalar) and adding the result to another row (column).

**Application 1. Computing determinants:** see the seminars.

We remind that these operations can be used for computing determinants and none of them turns a non-zero determinant into a zero and, also, none of them turns a zero determinant into a non-zero one. One can easily deduce from this fact that performing these operations on a matrix $A$ we get a matrix $B$ for which maximum number of columns (rows) we can choose from the columns (rows) of such that none of them is a linear combination of the others is the same, hence $A$ and $B$ have the same rank.

This remark can be used for developing another algorithm for finding the rank of a matrix.

**Application 2. Computing the rank of a matrix:**

For finding the rank of a matrix $A \in M_{m,n}(K)$ we may proceed by performing certain elementary operations on the rows and columns of $A$ and on the matrices that we successively get this way. Our purpose is to transform $A$ (by using elementary operations) into a matrix $B$ for which the only non-zero elements are on the main diagonal, preferably at the beginning (if such elements exist). This way, from $A$ we get a matrix

$$B = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & a_{22} & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{rr} & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{pmatrix}$$

with all $a_{11}$, $a_{22}, \dots$, $a_{rr}$ non-zero. The matrices $A$ and $B$ have the same rank which is $r$.

**Examples** - see the seminar

**Application 3. Solving systems of linear equations** by using **Gauss elimination** algorithm. Let $K$ be a field and let us consider the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1}$$

over $K$ with tha augmented matrix $\overline{A}$. This algorithm is based on the fact that

(i) interchanging of two equations of (1),

(ii) multiplying an equation of (1) by a non-zero element $\alpha \in K$,

(iii) multiplying an equation of (1) by $\alpha \in K$ and adding the resulted equation to another one, are operations which lead us to systems whic are equivalent to (1). Since all these operations act on the coefficients and constant terms of the system, it is quite obvious that these operations can be performed as elementary row operations on the system augmented matrix.

Thus, we can infer that providing elementary row operations on the augmented matrix of (1), we get tha augmented matrix of an equivalent system. **Gaussian elimination** (also known as **row reduction**) is an algorithm which uses row elementary operations on some matrices resulted from $\overline{A}$ in order to get a matrix with a number zero entries at the beginning of each row which strictly increases while we descend in the matrix (matrix known as **echelon matrix** or **echelon form**). This procedure corresponds to a partial elimination of some unknowns to get an equivalent system which can be easier solved.

**Definition 15.** A matrix $A \in M_{mn}(K)$ is in an **echelon form** with $k \geq 1$ non-zero rows if:

(1) the rows $1, \dots, k$ are non-zero and the rows $k+1, \dots, m$ consists only of 0;

(2) if $N(i)$ is the number of zeros at the beginning of the row $i$ ($i \in \{1, \dots, k\}$), then

$$0 \leq N(1) < N(2) < \cdots < N(k)\,.$$

A $k$ non-zero rows echelon form with $N(1) = 0$, $N(2) = 1$, $N(3) = 2, \dots$, $N(k) = k - 1$ is called **trapezoidal form**.

**Remarks 16.** a) Any matrix can be brought to an echelon form by elementary row operations.

b) A square matrix of size $n$ is invertible if and only it can be brought to a trapezoidal form with $n$ non-zero rows by using only elementary row operations (such a matrix is called **triangular form (matrix)**).

c) A square matrix of size $n$ is invertible if and only it can be brought to the identity matrix $I_n$ by using only elementary row operations.

Thus, the purpose of Gauss elimination is to successively use elementary operations on the rows of the augmented matrix $\overline{A}$ of (1) in order to bring it to an echelon form $B$. If we manage to do this, then $B$ is the augmented matrix of an equivalent system. In some forms of Gauss elimination, and we plan to use this form, the purpose is to bring $\overline{A}$ to a trapezoidal form

$$B = \begin{pmatrix} a'_{11} & a'_{12} & a'_{13} & \dots & a'_{1k} & a'_{1,k+1} & \dots & a'_{1n} & b'_1 \\ 0 & a'_{22} & a'_{23} & \dots & a'_{2k} & a'_{2,k+1} & \dots & a'_{2n} & b'_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a'_{kk} & a'_{k,k+1} & \dots & a'_{kn} & b'_k \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Some information on the given system can be easily read from this form. E.g. the rank of $\overline{A}$ is (the rank of $B$ which is) the number of the nonzero elements on the diagonal of $B$ and these nonzero elements on the diagonal of $B$ provide us with the main unknowns.

**Remarks 17.** a) Finding a trapezoidal form is not always possible by using only row elementary operations. Sometimes, we have to interchange two columns of the first $n$ columns, hence columns corresponding to the matrix of a certain equivalent system. This is, obviously, allowed since this means that we commute the two corresponding terms in each equation of this system and this is possible based on the commutativity of the addition in $K$.
b) If, during this algorithm, one finds a row for which all the elements are 0, except for the last one, which is $a \in K^*$, then (1) is inconsistent since it is equivalent to a system which contains the equality $0 = a$ which is not possible.

Assume that we brought $A$ to the above mentioned trapezoidal form $B$. This means that (1) is equivalent to a system of the form

$$\begin{cases} a'_{11}x_1 + a'_{12}x_2 + \dots + a'_{1,k-1}x_{k-1} + a'_{1k}x_k + a'_{1,k+1}x_{k+1} + \dots + a'_{1n}x_n = b'_1 \\ \qquad a'_{22}x_2 + \dots + a'_{2,k-1}x_{k-1} + a'_{2k}x_k + a'_{2,k+1}x_{k+1} + \dots + a'_{2n}x_n = b'_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \qquad\qquad a'_{k-1,k-1}x_{k-1} + a'_{k-1,k}x_k + a'_{k-1,k+1}x_{k+1} + \dots + a'_{k-1,n}x_n = b'_{k-1} \\ \qquad\qquad\qquad a'_{kk}x_k + a'_{k,k+1}x_{k+1} + \dots + a'_{kn}x_n = b'_k \end{cases}$$

(possibly with the unknowns succeeding in a different way, if we permuted columns) The main unknowns $x_1, \dots, x_k$ can be easily computed starting from the last equation of this system.

**Remarks 18.** a) A few more steps in Gauss elimination allow us to bring $\overline{A}$ by elementary row operations and, if necessary, by switching columns different from the last one to the following trapezoidal form

$$B = \begin{pmatrix} a''_{11} & 0 & 0 & \dots & 0 & a''_{1,k+1} & \dots & a''_{1n} & b''_1 \\ 0 & a''_{22} & 0 & \dots & 0 & a''_{2,k+1} & \dots & a''_{2n} & b''_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a''_{kk} & a''_{k,k+1} & \dots & a''_{kn} & b''_k \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

with $a''_{11}, a''_{22}, \dots, a''_{kk}$ non-zero (of course, this is possible only if the system (1) is consistent, otherwise, some non-zero elements may appear in the last column, bellow $b''_k$). One can easily notice the advantage we have when we form the equivalent system of (1) provided by $B$. This algorithm is known as **Gauss-Jordan elimination**.

b) Moreover, we can bring the augmented matrix of a consistent system to the following form:

$$B = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & a'''_{1,k+1} & \dots & a'''_{1n} & b'''_1 \\ 0 & 1 & 0 & \dots & 0 & a'''_{2,k+1} & \dots & a'''_{2n} & b'''_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & a'''_{k,k+1} & \dots & a'''_{kn} & b'''_k \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Now, it is straightforward to express the main unknowns as linear combinations of the side unknowns.

**Application 4. Computing the inverse of a matrix:** Let $K$ be a field, $n \in \mathbb{N}^*$ and let us consider $A = (a_{ij}) \in M_n(K)$ a matrix with $d = \det A \neq 0$. We rimind that the matrix equation

$$A \begin{pmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix} = \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} \tag{2}$$

is an equivalent form of a (consistent) Cramer system and that its unique solution is

$$\begin{pmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix} = A^{-1} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix}.$$

Let us take $j = 1$ and $\begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Then $\begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix}$ is the first column of the matrix $A^{-1}$, i.e.

$$\begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} = A^{-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} d^{-1}\alpha_{11} \\ d^{-1}\alpha_{12} \\ \vdots \\ d^{-1}\alpha_{1n} \end{pmatrix}$$

(we remind that in our previous courses we denoted by $\alpha_{ij}$ the cofactor of $a_{ij}$). Of course,

$$\begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} = I_n \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix}$$

By means of Gauss-Jordan algorithm, one deduces that the augmented matrix of the system (2) can be brought by elementary row operations to the following form

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & d^{-1}\alpha_{11} \\ 0 & 1 & 0 & \dots & 0 & d^{-1}\alpha_{12} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & d^{-1}\alpha_{1n} \end{pmatrix}.$$

Taking, successively, $j = 2$ and $\begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, then $j = 3$ and $\begin{pmatrix} b_{13} \\ b_{23} \\ \vdots \\ b_{n3} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$,..., $j = n$ and $\begin{pmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$, we form the corresponding systems (2) and we use the Gauss-Jordan algorithm to solve them. We perform exactly the same elementary operations as in the case $j = 1$ on the rows of each augmented matrix of a resulted system in order to bring the system matrix to the form $I_n$. We get:

$$\begin{pmatrix} 1 & 0 & \dots & 0 & d^{-1}\alpha_{21} \\ 0 & 1 & \dots & 0 & d^{-1}\alpha_{22} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & d^{-1}\alpha_{2n} \end{pmatrix}, \begin{pmatrix} 1 & 0 & \dots & 0 & d^{-1}\alpha_{31} \\ 0 & 1 & \dots & 0 & d^{-1}\alpha_{32} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & d^{-1}\alpha_{3n} \end{pmatrix}, \dots, \begin{pmatrix} 1 & 0 & \dots & 0 & d^{-1}\alpha_{n1} \\ 0 & 1 & \dots & 0 & d^{-1}\alpha_{n2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & d^{-1}\alpha_{nn} \end{pmatrix},$$

respectively. The constant terms column and, consequently, the solution of each system we solved is the column 2 of $A^{-1}$, column 3 of $A^{-1}$,..., column $n$ of $A^{-1}$, respectively.

Since we performed the same row operations on each of the previously mentioned $n$ systems, we can solve all of them using the same algorithm. This way one can finds an algorithm for computing the inverse of the matrix $A$: we start from the $n \times 2n$ matrix resulted by attaching the matrices $A$ and $I_n$

$$(A \mid I_n) = \left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right) \in M_{n,2n}(K)$$

and we perform successive elementary row operations (and only row operations) on this matrix and on the matrices successively resulted from this in order to transform the left size block into $I_n$. Remark 8 c) of the previous course ensures us that this is possible (if and only if $A$ is invertible). The resulted matrix is:

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & d^{-1}\alpha_{11} & d^{-1}\alpha_{21} & \dots & d^{-1}\alpha_{n1} \\ 0 & 1 & \dots & 0 & d^{-1}\alpha_{12} & d^{-1}\alpha_{22} & \dots & d^{-1}\alpha_{n2} \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & d^{-1}\alpha_{1n} & d^{-1}\alpha_{2n} & \dots & d^{-1}\alpha_{nn} \end{array}\right) = (I_n \mid A^{-1})$$

Thus, the right side block of the resulted matrix is the exactly the inverse matrix of $A$.

**Definition 19.** A square matrix resulted from the identilastty matrix after performing only one elementary operation is called **elementary matrix**.

**Remarks 20. (and examples ...)**

a) The elementary matrices resulted by switching rows (columns):

$$\begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ last0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

have the determinant $-1$.

b) The elementary matrices resulted by multiplying a row (column) with $\alpha \in K^*$:

$$\begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & \alpha & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

have the determinant $\alpha$.
c) The elementary matrices resulted by multiplying a row (column) by $\alpha \in K$ and adding the result to another row (column):

$$\begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & \alpha & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

have the determinant 1.

For the elementary matrices resulted by switching rows (columns), we have:

$$\begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

For the elementary matrices resulted by multiplying a row (column) with $\alpha \in K^*$, we have:

$$\begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & \alpha & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & \alpha^{-1} & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

For the elementary matrices resulted by multiplying a row (column) by $\alpha \in K$ and adding the result to another row (column), we have:

$$\begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & \alpha & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & -\alpha & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

Therefore, we can state the following:

**Lemma 21.** The inverse of an elementary matrix is also an elementary matrix.

**Lemma 22.** Let $m, n \in \mathbb{N}^*$. Any elementary operation on a matrix $A = (a_{ij}) \in M_{m,n}(K)$ is the result of the multiplication of $A$ with an elementary matrix. More precisely, any elementary operation on the rows (columns) of $A$ results by multiplying $A$ on the left (right) side with the elementary matrix resulted by performing the same elementary operation on $I_m$ ($I_n$, respectively).

*Proof.* We check this property for rows. For columns — HOMEWORK.

Let us switch the rows $i$ and $j$ of $I_m$ and let us multiply the resulted elementary matrix with $A$. The matrix

$$\begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ii} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

is

$$\begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ii} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

which is exactly the matrix resulted from $A$ by switching the rows $i$ and $j$.

Let $\alpha \in K^*$, let us multiply the $i$'th row of $I_m$ by $\alpha$ and let us multiply the resulted elementary matrix with $A$. The matrix

$$\begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & \alpha & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ii} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

is

$$\begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ \alpha a_{i1} & \dots & \alpha a_{ii} & \dots & \alpha a_{ij} & \dots & \alpha a_{in} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

which is exactly the matrix resulted from $A$ by multiplying the $i$'th row by $\alpha$.

Let $\alpha \in K$, let us take the elementary matrix that we get from $I_m$ after multiplying the $j$'th row by $\alpha$ and adding the result to the $i$'th row, and let us multiply this elementary matrix with $A$. The matrix

$$\begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & \alpha & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ii} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

is

$$\begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1} + \alpha a_{j1} & \dots & a_{ii} + \alpha a_{ji} & \dots & a_{ij} + \alpha a_{jj} & \dots & a_{in} + \alpha a_{jn} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

which is exactly the matrix that we get from $A$ after multiplying the $j$'th row by $\alpha$ and adding the result to the $i$'th row. □

**Corollary 23.** Any invertible matrix is a product of elementary matrices.

**Theorem 24.** Let $n \in \mathbb{N}^*$. For any matrices $A, B \in M_n(K)$ we have $\det(AB) = \det A \cdot \det B$.

*Proof.* □