COURSE 6

The rank of a matrix

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

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

$$\begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \dots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \dots & a_{i_2 j_k} \\ \vdots & \vdots & & \vdots \\ a_{i_l j_1} & a_{i_l j_2} & \dots & a_{i_l j_k} \end{pmatrix}$$

formed by taking the elementels of A which are situated at the intersections of the rows $i_1, \dots, i_k$ with the columns $j_1, \dots, j_l$ 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) $\operatorname{rank} A \le \min\{m, n\}$.
b) If $A \in M_n(K)$ then $\operatorname{rank} A = n$ dif and only if $\det A \neq 0$.
c) $\operatorname{rank} A = \operatorname{rank} {}^t A$.

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

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

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

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

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

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

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

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

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 $1 \le j_1 < \dots < j_r \le p$. (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:

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

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 $(\alpha_1, \dots, \alpha_n) \in K^n$ is a **solution of the system** (1) if the all the equalities resulted by replacing x_i with α_i $(i = 1, \dots, n)$ 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 $(0, 0, \dots, 0) \in K^n$ 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-Cappelli) 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. $\square$

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 $x_{r+1},\dots,x_n$ are side unknowns. We can assign them arbitrary "values" from K, $\alpha_{r+1},\dots,\alpha_n$, 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).