Topic 1: Matrices and Systems of Linear Equations.

Let us start with a review of some linear algebra concepts we have already learned, such as matrices, determinants, etc. Also, we shall review the method of solving systems of linear equations, as well as the geometric interpretation of the solution set.

1. Matrices and Determinants. Rank of a matrix. Matrix multiplication and inverse of a matrix

An $n \times m$ matrix is a set of $n \times m$ real numbers arranged in n rows and m columns. A matrix A of order $n \times m$ is represented as follows

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

that is the element a_{ij} is the entry in the i'th row and the j'th column. We write $A \in M_{n\times m}$. Sometimes, we denote the matrix that has elements a_{ij} using an older convention: $A = (a_{ij})_{j=1,\ldots,m}^{i=1,\ldots,n}$ or $A = (a_{ij})_{ij}$.

The elements a_{ii} are called the **diagonal** of A.

When a matrix has n rows and m columns we say that the matrix is of the dimension (size) $n \times m$. If the number of rows and columns is the same (i.e., if $m = n$), the matrix is called a **square matrix**.

The square matrix

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

is known as the identity matrix of dimension n, and is denoted by I_n.

Definition. Given a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \in M_{n\times m}$$

we define the **transpose matrix** $A^t \in M_{m\times n}$ (or A^*) of A, as the matrix whose row i is equal to column i of A. That is,

$$A^t = A^* = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix} \in M_{m\times n}$$

Definition. Given a square matrix,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in M_{n\times m}$$

the **trace** of A is the following real number

$$\text{trace}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

1.1. **Addition and multiplication by scalars.** Matrices of the same dimension can be **added**. The sum is obtained by adding the elements that are located in the same row and column of the two matrices. Thus, if $A = (a_{ij})_{j=1,\ldots,m}^{i=1,\ldots,n}$ and $B = (b_{ij})_{j=1,\ldots,m}^{i=1,\ldots,n}$ (as you see the matrices have the same dimension), then

$$A + B = (a_{ij} + b_{ij})_{j=1,\ldots,m}^{i=1,\ldots,n}$$

Example.

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

$$A + B = \begin{pmatrix} 2+1 & 1+4 & 3+0 \\ 9+5 & 6+(-2) & 5+(-3) \end{pmatrix} = \begin{pmatrix} 3 & 5 & 3 \\ 14 & 4 & 2 \end{pmatrix}$$

Multiplication of a matrix by a real number is defined similarly. If $A = (a_{ij})_{j=1,\ldots,m}^{i=1,\ldots,n}$ and $\lambda \in \mathbb{R}$, then

$$\lambda A = (\lambda a_{ij})_{j=1,\ldots,m}^{i=1,\ldots,n}$$

Example. Let λ be any (real) number and consider the matrix

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 9 & 6 & 5 \end{pmatrix}$$

then

$$\lambda A = \begin{pmatrix} \lambda\cdot 2 & \lambda\cdot 1 & \lambda\cdot 3 \\ \lambda\cdot 9 & \lambda\cdot 6 & \lambda\cdot 5 \end{pmatrix}$$

If we fix $\lambda = 7$

$$7A = \begin{pmatrix} 7\cdot 2 & 7\cdot 1 & 7\cdot 3 \\ 7\cdot 9 & 7\cdot 6 & 7\cdot 5 \end{pmatrix} = \begin{pmatrix} 14 & 7 & 21 \\ 63 & 42 & 35 \end{pmatrix}$$

The matrix addition and multiplication by scalars satisfy the following properties that can be easily derived from the above definitions:

Proposition 1. Let A, B and C be two matrices of the same dimension and let α and β be any real numbers. Then,

(1) $A + B = B + A$ (commutativity).
(2) $A + (B + C) = (A + B) + C$ (associativity).
(3) $\alpha(A + B) = \alpha A + \alpha B$.
(4) $(\alpha + \beta)A = \alpha A + \beta A$.
(5) $\alpha(\beta A) = (\alpha\beta)A$.
(6) $(A + B)^t = A^t + B^t$.

1.2. **Matrix Multiplication.** Consider now $A = (a_{ij})_{j=1,\ldots,m}^{i=1,\ldots n}$ an $n \times m$ matrix and $B = (b_{ij})$ of size $m \times l$. That is, the number of columns of matrix A equals the number of rows in matrix B. This is a necessary (and sufficient) condition for being able to calculate the product matrix $A \cdot B$.

Definition. The product matrix $C = A \cdot B$ is the matrix with n rows (same number as A) and l columns (same as B) such that the element

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}$$

that is, the element of the product matrix in the position i, j is obtained by multiplying row i of A times column j of B.

Example. Consider the matrices

$$A = \begin{pmatrix} 2 & 1 & 5 \\ -3 & 0 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 6 \\ 7 & -4 \\ 8 & 0 \end{pmatrix}$$

the matrix $C = A \cdot B$ is the 2×2 matrix

$$C = \begin{pmatrix} 49 & 8 = 2 \cdot 6 + 1 \cdot (-4) + 5 \cdot 0 \\ 13 & -18 \end{pmatrix}$$

Remark. In general, the product of matrices is not commutative. In fact, in many cases it is only possible to calculate one product, but not the other.

For instance, if we consider the matrices

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 & 5 \\ 1 & 4 & 4 \end{pmatrix}$$

we can calculate $A \cdot B$. But, $B \cdot A$ is not defined, since the number of columns of B does not coincide with the number of rows of A.

If we take

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$$

then,

$$A \cdot B = \begin{pmatrix} 7 & 3 \\ 8 & 3 \end{pmatrix} \qquad B \cdot A = \begin{pmatrix} 5 & 7 \\ 4 & 5 \end{pmatrix}$$

Proposition 2. If the product AB exists, then the product $B^t A^t$ is also possible and

$$(AB)^t = B^t A^t$$

1.3. **Equivalent Matrices. The Gauss-Jordan elimination Method.** Given a matrix A we say that B **is equivalent to** A if we can transform A into B using any combination of the following elementary operations:

- Multiply a row of A by any real number not equal to zero.
- Interchange two rows.
- Add a row of A to any other row.

These three operations can be described by means of matrix products

- Multiplying the i^{th} row of an $n \times m$ matrix by a real number a is equivalent to multiplying on the left by the identity matrix I_n in which the element in the position ii is replaced by a.

For instance, if we consider a matrix

$$\begin{pmatrix} 2 & 4 \\ 3 & 6 \\ 7 & 9 \end{pmatrix}$$

and want to multiply the 2nd row by 5 we have to multiply this matrix by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- To swap two rows, say rows i and j, the only thing we have to do is to multiply on the left by an identity matrix with rows i and j swapped
  For instance, if in the previous matrix:

$$\begin{pmatrix} 2 & 4 \\ 3 & 6 \\ 7 & 9 \end{pmatrix}$$

  we want to exchange rows 1 and 3, then we multiply

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & 4 \\ 3 & 6 \\ 7 & 9 \end{pmatrix}$$

- The operation of adding the j^{th} row multiplied by a real number a, to the i^{th} row is equivalent to multiply on the left by the identity matrix but with an a in the entry ij.
  For instance, if in the matrix

$$\begin{pmatrix} 2 & 4 \\ 3 & 6 \\ 7 & 9 \end{pmatrix}$$

  we want to add to row 2 seven times row 3, we can do it in the following way,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 4 \\ 3 & 6 \\ 7 & 9 \end{pmatrix}$$

  that is, we put a 7 into the position $2, 3$

Definition 3. We say that a matrix $A \in M_{n\times m}$ is in (row) **echelon form** if it satisfies the following:

(1) All zero rows are at the bottom of the matrix.
(2) For each row $i = 1, \ldots, n$, if the first non-zero element is a_{ij} (that is, $a_{ij} \neq 0$ and $a_{ik} = 0$ for every $k < j$), then $a_{l,k} = 0$ for every $l > i$ and every $k \leq l$.

It follows that, if the matrix A is in echelon form, then given any row i, if its first element that is not equal to zero is a_{ij} then $a_{kl} = 0$ whenever $k > i$ and $l \leq j$. In an echelon matrix, every row (except perhaps the first one) starts with 0 and each row $i + 1$ starts with at least one more 0 than the preceding row i.

For example,

$$\begin{pmatrix} 5 & 3 & 0 & 4 & 1 \\ 0 & 1 & 6 & 2 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

are matrices in echelon form. While,

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

is not.

The Gauss-Jordan elimination method gives us a systematic way of obtaining an echelon form for any matrix by means of elementary matrix operations. We will explain the method by means of an example. Consider the following matrix,

$$\begin{pmatrix} 0 & 3 & 2 & 5 & 7 \\ 1 & 7 & 2 & 4 & 3 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 4 & 7 \end{pmatrix}$$

(1) Reorder the rows so that all the rows with all elements equal to zero, if any, are below all other rows.

$$\begin{pmatrix} 0 & 3 & 2 & 5 & 7 \\ 1 & 7 & 2 & 4 & 3 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 5 & 0 & 4 & 7 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(2) Look for the first column that doesn't have all zeros.

$$\begin{pmatrix} \color{red}{0} & 3 & 2 & 5 & 7 \\ \color{red}{1} & 7 & 2 & 4 & 3 \\ \color{red}{0} & 0 & 0 & 1 & 3 \\ \color{red}{0} & 5 & 0 & 4 & 7 \\ \color{red}{0} & 0 & 0 & 0 & 0 \end{pmatrix}$$

(3) Reorder the rows again, so that all the zeros in this column are below all the non-zero elements.

$$\begin{pmatrix} \color{red}{1} & 7 & 2 & 4 & 3 \\ \color{red}{0} & 3 & 2 & 5 & 7 \\ \color{red}{0} & 0 & 0 & 1 & 3 \\ \color{red}{0} & 5 & 0 & 4 & 7 \\ \color{red}{0} & 0 & 0 & 0 & 0 \end{pmatrix}$$

(4) If the matrix is already in echelon form, we are finished.

(5) If not, we look for the first element that violates the echelon condition (we call it **pivotal**) and from now on forget all the rows above it.

$$\begin{pmatrix} 1 & 7 & 2 & 4 & 3 \\ 0 & 3 & 2 & 5 & 7 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 5 & 0 & 4 & 7 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(6) Repeat the previous step (ignoring the upper rows that have been already dealt with). If the matrix is already in echelon form, we are done.

$$\begin{pmatrix} 1 & 7 & 2 & 4 & 3 \\ 0 & 3 & 2 & 5 & 7 \\ 0 & 5 & 0 & 4 & 7 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(7) If not, eliminate all the non-zero numbers that are in the same column as the pivotal element (let it be a_{ij}) and below it by adding multiples of row i to the multiples of rows below it:

$$\begin{array}{r} \\ \\ 3\cdot\text{row }3 - 5\cdot\text{row }2 \\ \\ \\ \end{array} \begin{pmatrix} 1 & 7 & 2 & 4 & 3 \\ 0 & 3 & 2 & 5 & 7 \\ 0 & 15-15 & 0-10 & 12-25 & 21-35 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 7 & 2 & 4 & 3 \\ 0 & 3 & 2 & 5 & 7 \\ 0 & 0 & -10 & -13 & -14 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(8) If the matrix is already in echelon form, we are done. Otherwise, repeat the operations with the rows that are still not echelon form, until you are done.

1.4. **Determinants.** To any square matrix one can assign a real number called the **determinant** of the matrix. As we are going to see, this number is important and useful. We shall define it by induction.

Let A be a 1×1 matrix, i.e. it has a single row and column. In other words, A is a scalar $A = (a)$. The determinant in this case is simply defined to be a.

Let A be a 2×2 matrix. Then, the determinant is defined as

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb$$

Let A be a 3×3 matrix. In this case we have two equivalent ways of defining the determinant:

(1) Sarrus rule:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{31}a_{22}a_{13}-a_{32}a_{23}a_{11}-a_{33}a_{21}a_{12}$$

(2) Expanding by a row (or column):

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

We don't have to be using the first row or column; any row or column would work. We have to be careful, though, with signs, since the term multiplied also by the element a_{ij} has to be multiplied by $(-1)^{i+j}$.

Example. If we want to expand the determinant by column 2,

$$\begin{vmatrix} 1 & 2 & 1 \\ 4 & 3 & 5 \\ 3 & 1 & 3 \end{vmatrix} = (-1)^{1+2}2\begin{vmatrix} 4 & 5 \\ 3 & 3 \end{vmatrix} + (-1)^{2+2}3\begin{vmatrix} 1 & 1 \\ 3 & 3 \end{vmatrix} + (-1)^{2+3}1\begin{vmatrix} 1 & 1 \\ 4 & 5 \end{vmatrix} =$$

$$-2\cdot(-3)+3\cdot(0)-(1)\cdot 1 = 5$$

For matrices of higher dimensions the rule is the same as for the 3×3 matrices: we can expand by rows or columns in such a way that the determinant of a 4×4 matrix can be calculated by calculating 4 determinants of 3×3 matrices.

In general, if $A=(a_{ij})$ is a square $n\times n$ matrix , the **minor** complementary to the element a_{ij} of A is the determinant of the $n-1\times n-1$ submatrix that is obtained by eliminating the row i and the column j from the matrix A. The **cofactor** A_{ij} of the element a_{ij} of A is the minor complementary to a_{ij} multiplied by the factor $(-1)^{i+j}$. With this notation, we can write the expansion of the determinant of a matrix A by row i as,

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{im}A_{im}$$

or by column j,

$$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}$$

Example. Consider the determinant

$$\begin{vmatrix} 1 & 2 & 0 & 3 \\ 4 & 7 & 1 & 1 \\ 1 & 3 & 3 & 1 \\ 0 & 2 & 0 & 7 \end{vmatrix}$$

It is best to expand this determinant by row 4 or column 3. This is so, because there are more zeroes in them. So that in the end, we would have to do fewer calculations. Let us expand by column 3:

$$\begin{vmatrix} 1 & 2 & 0 & 3 \\ 4 & 7 & 2 & 1 \\ 1 & 3 & 3 & 1 \\ 0 & 2 & 0 & 7 \end{vmatrix} = 0\begin{vmatrix} 4 & 7 & 1 \\ 1 & 3 & 1 \\ 0 & 2 & 7 \end{vmatrix} + (-1)^{3+2}2\begin{vmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \\ 0 & 2 & 7 \end{vmatrix} + (-1)^{3+3}3\begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & 1 \\ 0 & 2 & 7 \end{vmatrix} + 0\begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & 1 \\ 1 & 3 & 1 \end{vmatrix}$$

As you see from this example, the more zeroes there are, the easier it is to do the calculations. We shall now explain how to make there as many zeroes as possible.

Proposition 4.

(1) Let A and B be square matrices of the same size. Then,
$$\det(A \cdot B) = \det(A) \cdot \det(B)$$
(2) If all elements of a row or a column of a matrix are equal to zero, then the determinant also equals to zero.

The remaining properties are easily derived from those above and can be used to increase the number of zeros.

Proposition 5.

(1) If we multiply a row or a column of a matrix by a number, then the determinant is also multiplied by the same number.
(2) If we swap two rows (or two columns) the determinant changes sign.
(3) If we add to a row (or a column) a multiple of another row (respectively, column) the determinant does not change.
(4) If two rows or columns are equal, then the determinant is equal to 0.

Using the rules above, we can simplify the computation of determinants by using the same elementary operations we have used to put the matrix into its echelon form.

Example. Consider the determinant:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 5 & 1 \\ 3 & 1 & 4 & 2 \\ 2 & 5 & 5 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -1 & -7 \\ 0 & -5 & -5 & -10 \\ 0 & 1 & -1 & -7 \end{vmatrix} =$$
$$= \begin{vmatrix} -1 & -1 & -7 \\ -5 & -5 & -10 \\ 1 & -1 & -7 \end{vmatrix} = - \begin{vmatrix} 1 & 1 & 7 \\ -5 & -5 & -10 \\ 1 & -1 & -7 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 7 \\ 0 & 0 & 25 \\ 0 & -2 & -14 \end{vmatrix} =$$
$$= - \begin{vmatrix} 1 & 1 & 7 \\ 0 & -2 & -14 \\ 0 & 0 & 25 \end{vmatrix} = 50$$

1.5. **Rank of a Matrix.**

Remark. The echelon form obtained by the above procedure is not unique. That is, when we apply the method above to reduce a matrix to an echelon form, the final matrix that we obtain, depends on the exact way the steps are carried out. However, one can prove that the number of zero rows obtained is independent of the procedure to find the echelon form.

Definition 6. Given any matrix A we define **rank of** A to be the number of rows not all equal to zero that the matrix has in any of its echelon forms.

Example. Consider the matrix
$$A = \begin{pmatrix} 0 & 3 & 2 & 5 & 7 \\ 1 & 7 & 2 & 4 & 3 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 4 & 7 \end{pmatrix}$$

it is equivalent to the echelon matrix:

$$= \begin{pmatrix} 1 & 7 & 2 & 4 & 3 \\ 0 & 3 & 2 & 5 & 7 \\ 0 & 0 & -10 & -13 & -14 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

so that the rank of A is four.

There is an equivalent definition of the matrix rank that uses determinants.

Proposition 7. Let A be any matrix. The rank of A is the dimension of the largest square matrix with a non-zero determinant that we can construct from the original matrix by eliminating rows and columns.

Example.

(1) Take a matrix

$$\begin{pmatrix} 1 & 2 & 5 \\ 2 & 4 & 9 \end{pmatrix}$$

The rank of this matrix can't be more than 2 since it is impossible to construct inside it a 3×3 matrix. Let us see if we can construct a 2x2 matrix with a non-zero determinant

$$\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0$$

doesn't work

$$\begin{vmatrix} 1 & 5 \\ 2 & 9 \end{vmatrix} = -1 \neq 0$$

Since, this determinant does not equal to zero, the rank of this matrix is 2.

(2) If we now take

$$\begin{pmatrix} 1 & 2 & 5 \\ 2 & 4 & 10 \end{pmatrix}$$

then, as before, the rank can't be more than 2. Nor can it be smaller than 1, since for that the matrix would have to be a matrix of zeroes. Let us see if it is 2:

$$\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0, \begin{vmatrix} 1 & 5 \\ 2 & 10 \end{vmatrix} = 0, \begin{vmatrix} 2 & 5 \\ 4 & 10 \end{vmatrix} = 0$$

Therefore, the rank is 1.

Remark. The rank of two equivalent matrices is the same

1.6. **Inverse Matrix.** Given an $n \times n$ **square** matrix A, we say that it has an inverse if there exists an $n \times n$ matrix, which we denote by A^{-1}, such that $A^{-1}A = AA^{-1} = I_n$, where I_n is the identity matrix of order $n \times n$.

Remark. Is there a unique inverse matrix? In other words, is it possible that there are two different matrices, say B and C such that $BA = AB = I_n$ and $CA = AC = I_n$? Let us show that, if these equations hold then, we must have $B = C$.

Since, $CA = I_n$, multiplying by B on the right we get that

$$(CA)B = I_nB = B$$

and since,

$$(CA)B = C(AB) = CI_n = C$$

we see that $B = C$.

Proposition 8. A square matrix has an inverse if and only if its determinant is not equal to zero. Equivalently, an $n \times n$ square matrix has an inverse if and only if it has the rank n (that is, it has *full rank*)

Therefore, using the properties of determinants it is easy to derive the following:

Proposition 9. If A has an inverse, then $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

Proof: Since, $A \cdot A^{-1} = I_n$, we must have that $\det(A \cdot A^{-1}) = \det(I_n)$. One checks easily that $\det(I_n) = 1$. And, since $\det(A \cdot A^{-1}) = \det(A)\det(A^{-1})$ we must have that $\det(A)\det(A^{-1}) = 1$ and the proposition follows.

Proposition 10. Let A and B be $n \times n$ square matrices. Then, $A \cdot B$ and $B \cdot A$ have inverses only if A and B have inverses. Furthermore,

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

There are various methods of computing the inverse of a matrix, but undoubtedly the easiest is the Gauss-Jordan elimination method that we have used to put matrices in echelon form.

Consider the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Construct the new matrix

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right)$$

To this entire matrix we apply the *elementary operations* until the left side becomes an identity matrix. It can be shown that this can be done if and only if the original matrix has full-rank, in which case we obtain

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & 1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{array}\right)$$

and the matrix on the right is the inverse of A.

This method works since the matrix that we obtain on the right is a product of all those matrices by which we were multiplying the original matrix to turn it into the identity matrix.

Example. Consider the matrix

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

construct the larger matrix:

$$\begin{pmatrix} 1 & 1 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & 1 & | & 0 & 1 & 0 \\ 1 & 0 & 1 & | & 0 & 0 & 1 \end{pmatrix}$$

Make the following operations

$$(f_3 - f_1) \sim \begin{pmatrix} 1 & 1 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & 1 & | & 0 & 1 & 0 \\ 0 & -1 & 1 & | & -1 & 0 & 1 \end{pmatrix} (f_3 + f_2) \sim \begin{pmatrix} 1 & 1 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & 1 & | & 0 & 1 & 0 \\ 0 & 0 & 2 & | & -1 & 1 & 1 \end{pmatrix}$$

At this point we see that this is a rank 3 matrix, so it can be inverted

$$(2f_2 - f_3) \sim \begin{pmatrix} 1 & 1 & 0 & | & 1 & 0 & 0 \\ 0 & 2 & 0 & | & 1 & 1 & -1 \\ 0 & 0 & 2 & | & -1 & 1 & 1 \end{pmatrix} (2f_1 - f_2) \sim \begin{pmatrix} 2 & 0 & 0 & | & 1 & -1 & 1 \\ 0 & 2 & 0 & | & 1 & 1 & -1 \\ 0 & 0 & 2 & | & -1 & 1 & 1 \end{pmatrix}$$

Finally, dividing by 2

$$\sim \begin{pmatrix} 1 & 0 & 0 & | & 1/2 & -1/2 & 1/2 \\ 0 & 1 & 0 & | & 1/2 & 1/2 & -1/2 \\ 0 & 0 & 1 & | & -1/2 & 1/2 & 1/2 \end{pmatrix}$$

Therefore, the inverse matrix is

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

For those who are absolutely in love with formulas, there does exist a formula that provides and inverse of a matrix. For a square matrix A, its adjoint matrix ($\mathrm{Adj}(A)$) is the matrix of cofactors of A. That is, an element in the i^{th} row and j^{th} column of $\mathrm{Adj}(A)$, is A_{ij}.

Proposition. *If* $|A| \neq 0$*, then*

$$A^{-1} = \frac{1}{|A|} \left(\mathrm{Adj}(A)\right)^t$$

2. Systems of Linear Equations

A system of linear equations is a system of equations of the form

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = & b_1 \\ \vdots & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = & b_m \end{cases}$$

where a_{ij} and b_k are known real numbers and $x_1, \ldots, x_n$ are the unknowns (or variables) of the system.

A system of linear equations can be rewritten using matrix notation as:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

A solution of the system is an $n-$dimensional vector that solves the matrix equation. Or, in other words, it is the vector of real numbers $(x_1^*, \dots x_n^*)$ which satisfy all the equations of the system.

Definition 11. A system of linear equations is called **consistent** if it has a solution, otherwise it is called **inconsistent** or overdetermined. If the system of linear equations has multiple (actually, infinitely many) solutions, it is called **underdetermined**.

Example. The system of linear equations

$$\begin{cases} 2x + y = 5 \\ 4x + 2y = 7 \end{cases}$$

doesn't have a solution, so it is inconsistent

The system of equations

$$\begin{cases} x + y = 5 \\ 4y = 8 \end{cases}$$

has as its unique solution the vector $(3, 2)$. Thus, it is consistent and uniquely determined.

The system of equations

$$\begin{cases} x + y = 4 \\ 2x + 2y = 8 \end{cases}$$

has infinitely many solutions, since all vectors of the form $(x, 4 - x)$ solve it; therefore, it is consistent and underdetermined.

2.1. **The Rouché-Frobenius Theorem.** The Rouché-Frobenius theorem gives us a criterion to decide when the system is consistent or inconsistent. And, in the former case, how many parameters do we need to describe the set of solutions.

For instance, consider the system of equations

$$\begin{cases} x + y = 2 \\ 2x + 2y = 4 \end{cases}$$

It can be easily seen that the second equation is just the double of the first, and is, therefore, redundant. Thus, the set of solutions is the set of two-dimensional vectors (x, y) that satisfy the condition $y = 2 - x$. If we choose a value of x we, therefore, obtain the unique value of y that satisfies the system. This set,

$$\{(x, 2 - x) : x \in \mathbb{R}\}$$

can be described by means of a single parameter. We say that it has one degree of freedom.

The Rouché-Frobenius theorem tells us exactly how many parameters (degrees of freedom) we need to describe a solution of any system. Consider the system of linear equations,

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = & b_1 \\ \vdots & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = & b_m \end{cases}$$

with m equations and n unknowns. We define the **matrix of the system** to be

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

The **extended matrix** of the system is

$$(A|b) = \begin{pmatrix} a_{11} & \dots & a_{1n} & | & b_1 \\ \vdots & \dots & \vdots & | & \vdots \\ a_{m1} & \dots & a_{mn} & | & b_m \end{pmatrix}$$

Theorem 1 (Rouché-Frobenius)**.**

(1) The system is consistent if and only if $\operatorname{rank} A = \operatorname{rank}(A|b)$.

(2) Suppose the system is consistent (Hence, $\operatorname{rank} A = \operatorname{rank}(A|b) \leq n$). Then,

   (a) The system has a unique solution if and only if $\operatorname{rank} A = \operatorname{rank}(A|b) = n$.

   (b) The system is underdetermined if and only if $\operatorname{rank} A = \operatorname{rank}(A|b) < n$. In this case, the number of parameters necessary to describe the solutions of the system is $n - \operatorname{rank}(A)$.

Example. Consider the system

$$\begin{cases} x + y + 2z = 1 \\ 2x + y + 3z = 2 \\ 3x + 2y + 5z = 3 \end{cases}$$

Its extended matrix is

$$\begin{pmatrix} 1 & 1 & 2 & | & 1 \\ 2 & 1 & 3 & | & 2 \\ 3 & 2 & 5 & | & 3 \end{pmatrix}$$

Calculate the rank of A and of $(A|b)$:

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

Therefore, the rank of A and of $(A|b)$ is 2. Hence, the system is consistent and underdetermined.

Proposition 12. A homogeneous system (i.e., a system in which all free terms on the right hand side of the equation are equal to zero) is always consistent (It has at least one solution).

2.2. **Gauss-Jordan elimination method for solving systems of equations.** The Rouché-Frobenius theorem tells us when the system of linear equations has (or does not have) a solution. Combining it with the Gauss-Jordan elimination method we can explicitly obtain the solutions. The idea is that, if the two systems have equivalent extended matrices, they have the same solutions.

On the other hand, if a matrix is already put into its echelon form, it is very easy to calculate the solutions. Consider, for example, the system whose extended matrix is

$$\left(\begin{array}{ccc|c} 1 & 3 & 5 & 1 \\ 0 & 1 & 4 & 2 \\ 0 & 0 & 2 & 4 \end{array}\right)$$

The associated system has a unique solution. To find the solution we only have to recall that each column corresponds to a variable and that the column after the line corresponds to free terms. To give the solution, let us calculate the variables from the bottom up.

The last row tells us that $2z = 4$, so we can substitute $z = 2$ into the second equation and get $y + 8 = 2$, so that $y = -6$. Finally, we substitute these values into the first equation to obtain that $x - 18 + 10 = 1$. Hence, $x = 9$.

Let us now consider the system represented by the matrix:

$$\left(\begin{array}{ccc|c} 1 & 3 & 5 & 1 \\ 0 & 1 & 4 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

This system is consistent and underdetermined, since the rank of A and of the augmented matrix are both equal to 2 and the number of unknowns is 3. Therefore, we have one degree of freedom. The second equation is $y + 4z = 2$, which we rewrite as

$$y = 2 - 4z$$

Substituting now into the first equation, we obtain that $x + 3(2 - 4z) + 5z = 1$, meaning that

$$x = -5 + 7z$$

Therefore, the set of **all** solutions of our system is

$$\{(7z - 5, 2 - 4z, z) : z \in \mathbb{R}\}$$

Sometimes we have to be careful with the variables we leave as parameters. In the previous case any one of the variables would work fine. In the following example that's not the case.

$$\left(\begin{array}{ccc|c} 1 & 4 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Here the second equations tells us that $y = 1$. So, y cannot be left as a parameter. If, on the other hand we substitute now the value $y = 1$ into the first equation, we obtain that $x + 4 - z = 1$, so that $x = z - 3$. Here, the variables that can be left as parameters are x or z. The general solution of the system can be written as

$$\{(z - 3, 1, z) \in \mathbb{R}^3 : z \in \mathbb{R}\}$$

or, if you so prefer,

$$\{(x, 1, x + 3) \in \mathbb{R}^3 : x \in \mathbb{R}\}$$

The reason that systems represented with equivalent matrices have the same solutions are the following.

- If we multiply an equation by a real number different from zero, the solution doesn't change.
- If we reorder the equations, the solution of the system doesn't change
- If we add (a multiple of) one of the equations to another the solution doesn't change.

Thus, if two systems are represented by (row) equivalent matrices, their solutions are the same. The Gauss-Jordan elimination method consists of transforming a system into another one, representable by a matrix in echelon form. The solutions of the new system are easy to calculate and are the same as those of the original system.

2.3. **Cramer's Rule.** Suppose now we have a system of n equations in n unknowns. In this case the system can be represented by a square matrix. This system has a unique solution if and only if it has full-rank, which, in turn, is equivalent to having a non-zero determinant of the matrix. Cramer's rule provides a way of finding solutions of the system using the determinants. This is how it works.

Consider the system:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = & b_1 \\ \vdots & \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = & b_n \end{cases}$$

and suppose that

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \neq 0$$

so the system has a unique solution. Denote the unique solution of the system as $(x_1^*, x_2^*, \dots, x_n^*)$. Then,

$$x_1^* = \frac{\begin{vmatrix} b_1 & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \dots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}} \quad x_2^* = \frac{\begin{vmatrix} a_{11} & b_1 & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & b_n & \dots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}} \dots$$

$$\dots\ x_n^* = \frac{\begin{vmatrix} a_{11} & a_{12} & \dots & b_1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & b_n \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}}$$

We shall make two observations about his method.

- Cramer's method requires substantially more operations than the Gauss-Jordan elimination method.
- It is easy to extend Cramer's method to underdetermined systems (we won't do this in this class, though).