Rough Syllabus, Part 3.
Matrix arithmetics

1. Definition. A **matrix** is a rectangular array of numbers, symbols or expressions arranged in m rows and n columns. The individual items in a matrix are called the **elements** or **entries** of the matrix.

The general form of a matrix is:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

$a_{i,j}$ are the elements of the matrix, the first index of the elements is the **row index**, the second one is the **column index**. The dimension or size or type of a matrix is $(m \times n)$ (read m by n) if it has m rows and n columns.
Matrices are denoted usually by capital letters, and if it is necessary we mark their type: A, or $A_{(m\times n)}$.

Two matrices are equal, if they have the same dimension and the same elements on the corresponding positions.

In this part of discussion we consider matrices only with elements of real or complex numbers. In the followings we assume that matrices in the same formula have elements from the same numberfield, and constant multipliers are from there, as well.

2. Definition. The constant multiple of a matrix is a matrix of the same type,
$\alpha A_{(m\times n)} = B_{(m\times n)}$, where $b_{i,j} = \alpha a_{i,j}$

If we do not assume the special case that the entries of the matrix are numbers, then the constant multipliers must be from a number field such that the constant multiple of the entries are defined.

3. Definition. The sum of two matrices is defined if the matrices are from the same type (same dimension).
$A_{(m\times n)} + B_{(m\times n)} = C_{(m\times n)}$ where $c_{i,j} = a_{i,j} + b_{i,j}$

The addition obviously has the properties of the addition of the structure from where the elements are.
$A + B = B + A$, commutative;
$(A + B) + C = A(B + C)$ associative,
There is a neutral element for addition, the matrix with only 0 entries, the so called zero matrix;
Every matrix A has an inverse with respect to addition: $(-1)A$.

Remark: The matrices of the same type (of the same dimension) form a linear space. If e.g. the elements and constant multipliers are from the same number field, the dimension of this space is $m \cdot n$.

We can make the following observations, and remarks:
Matrices with one column or with one row have the same form what we usually use for vectors. Although we will see that vectors and matrices are far not the same, but calculations performed with vectors in matrix form give correct results. (A strong analogy can be seen from the previous remark.)

4. Definition. The product of two matrices is defined only if the number of column of the first one is equal to the number of the rows of the second one.

$$A_{(m\times l)}B_{(l\times n)} = C_{(m\times n)} \qquad \text{where} \qquad c_{i,j} = \sum_{k=1}^{l} a_{i,k}b_{k,j}.$$

Observe that the elements of the product are calculated like the dot product of a row from the first term and a column of the second term. (Later, discussing linear spaces, we will see that this is not the only way to calculate scalar product.)

The multiplication of matrices obviously not commutative, since the multiplication in the reversed order can not be performed if (applying the notations of the definition) $m \neq n$. If $m = n \neq l$, then the dimensions of AB and BA are not equal. Also in the case $n = m = l$ the product AB is usually not equal to BA.

In the case if $AB = BA$ then A and B **commute**.

Properties of multiplication:

- $(AB)C = A(BC)$ associative, if A, B, C can be multiplied in this order;
- $A(B + C) = AB + AC$ and $(B + C)G = BG + CG$ distributive with respect to matrix addition

Practical advise when multiplying matrices in a notebook: Arrange the terms of the product $AB = C$, in the following way:

$$\begin{array}{c|c} & B \\ \hline A & C \end{array}$$

Then there is less chance to misplace the elements of the result C, since the row and the column involved in an element of C are right in front of it in A and above it in B.

If 0 denotes a matrix who's elements are zeros, then $A_{(m\times l)}0_{(l\times n)} = 0_{(m\times n)}$, $0_{(m\times l)}B_{(l\times n)} = 0_{(m\times n)}$. (Warning: Usually $0A \neq A0$ because of their sizes.)

Observe that $AB = 0$ can happen when none of A and B are zero matrices.

Example:

$$A = \begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 2 & -1 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} -4 & -6 \\ -3 & -5 \\ 0 & 2 \\ 2 & 4 \end{bmatrix}, \qquad \begin{array}{cccc|cc} & & & & -4 & -6 \\ & & & & -3 & -5 \\ & & & & 0 & 2 \\ & & & & 2 & 4 \\ \hline 1 & 0 & -1 & 2 & 0 & 0 \\ 0 & 2 & -1 & 3 & 0 & 0 \end{array} \qquad AB = 0$$

Observe the followings:
If in the product $AB = C$ matrix B has only one column, then the result C has also one column, and it is the linear combination of the columns of A with constants in the only column in B.
If in the product $AB = C$ matrix A has only one row, then the result C has also one row, and it is the linear combination of the rows of B with constants in the only row in A.
A linear system of equations can be written in the form $A\mathbf{x} = \mathbf{b}$, where A is the coefficient matrix, and $\mathbf{x}$ is the vector of unknowns, and $\mathbf{b}$ is the vector on the right hand side.

Considering the scheme recommanded for notebook multiplication it is obvious that in the product $AB = C$
-every column of C is a linear combination of the columns of A with constant multipliers from the corresponding column of B;
-every row of C is a linear combination of the rows of B with constant multipliers from the corresponding row of A;
I.e. the columns of the product C are vectors such that the columns of B are multiplied by A one-by-one.
$AB = A[\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n] = [A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_n]$

The columns of a matrix can be considered as vectors, they are called the **column vectors** of the matrix. The same about rows.

5. Definition. i)
The **rank** of a matrix is the rank of the system of its column vectors. It is denoted by $\rho(A)$.

Since we have explained already with the properties of the row echelon form of the coefficient matrix of a linear system of equations that it is always equal to the rank of the row vectors:

Definition. ii)
The **rank** of a matrix is the rank of the system of its row vectors.

We will give a further "definition" of the rank of a matrix after discussing determinants.

1. Theorem. Considering the properties we have discussed about the rank of a system of vectors, the following inequalities are valid for the rank of a sum or of a product of matrices:
$\rho(A+B) \le \rho(A) + \rho(B), \quad \rho(A+B) \ge |\rho(A) - \rho(B)|, \quad \rho(AB) \le \min\big(\rho(A), \rho(B)\big).$
If $A_{(n\times k)}B_{(k\times m)} = C_{(n\times m)}$, then $\quad (k - \rho(C)) \le (k - \rho(A)) + (k - \rho(B)).$

6. Definition. The elements with the same row and column indices ($a_{i,i}$) are called **diagonal elements** and they form the **main diagonal** (or **diagonal**) of the matrix.

7. Definition. If in a matrix the number of rows is equal to the number of columns, then the matrix is called **quadratic** matrix.

8. Definition. A quadratic matrix who's every non-diagonal elements ($a_{i,j}, \quad i \ne j$) are zeros is called a **diagonal** matrix.
A quadratic matrix, who's every elements below the diagonal ($a_{i,j}, \quad i > j$) are zeros, is called an **upper triangular** matrix.
A quadratic matrix, who's every elements above the diagonal ($a_{i,j}, \quad i < j$) are zeros, is called a **lower triangular** matrix.
A diagonal matrix who's diagonal elements are all equal to 1, is called the **identity matrix** or **unit matrix**, and it is denoted by I.

Obviously diagonal matrices are upper triangle and lower triangle in the same time.

Some obvious facts:

- The sum and the product of diagonal matrices is a diagonal matrix.
- The sum and the product of triangular matrices is a triangular matrix.

Obviously that $A_{(m\times n)}I_{(n\times n)} = A(m\times n)$ and $I_{(m\times m)}A_{(m\times n)} = A_{(m\times n)}$, and only the identity matrix has the property that it does not changes any other matrix when multiplied by it.

9. Definition. The reflection of the matrix $A_(m \times n)$ with respect to its main diagonal, results in a matrix of type $(n \times m)$ which is called the **transposed** of A, and it is denoted by A^T. I.e. $a^{(T)}_{i,j} = a_{j,i}$.

10. Definition. Let $\overline{A}$ denote the matrix by taking the complex conjugate of the elements of A, then the matrix $\overline{(A^T)}$ is denoted by A^* and it is called the **adjoint** of A.
Warning: the expression "adjoint of A" is used also in an other meaning when talking about determinants.

2. Theorem. For transpose of matrices the followings are valid:
$(A^T)^T = A$, $(A+B)^T = A^T + B^T$, $(AB)^T = B^TA^T$, $\rho(A^T) = \rho(A)$.

Identity matrices behave like neutral elements with respect to multiplication, but for non-quadratic matrices we have to consider identity matrices with different sizes. So, when we want to investigate about multiplicative inverses of matrices, we have to talk about right inverse and left inverse because of the non-commutativity of multiplication, furthermore we should produce the different identity matrices. Altough there are different ways to define generalized inverses, in this semester we will talk only about the inverses of quadratic matrices.

Since matrix multiplication is not commutative, we have to look for such X and Y matrices for which $AX = I$ i.e. for a right inverse, and $YA = I$, i.e. for a left inverse.

3. Theorem. If A has a left and a right inverse, then they are equal. I.e. If $AX = YA = I$, then $X = Y$.

Proof: $Y = YI = Y(AX) = (YA)X = IX = X$.

Try to find an X for which $AX = I$.
$AX = A[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] = [A\mathbf{x}_1, A\mathbf{x}_2, \ldots, A\mathbf{x}_n] = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n,] = I$, where $\mathbf{e}_i$ is the i-th column of the identity matrix. Since two matrices are equal if their elements are equal,
$A\mathbf{x}_1 = \mathbf{e}_1$, $A\mathbf{x}_2 = \mathbf{e}_2$, $\ldots$, $A\mathbf{x}_n = \mathbf{e}_n$. This is a simultaneous system of systems of equations. If these equations can be solved, then n linearly independent vectors can be expressed as the linear combinations of the columns of A. Then A must have n linearly independent columns. So a matrix $A_{(n\times n)}$ can have an inverse only if its rank is n. When solving a system of equations the steps are depending only on the coefficient matrix, the right hand side only follows the steps. Solving the simultaneous systems one-by-one we should make the same steps n-times. We write the n right sides next to each others, and solve the systems simultaneously.

So, if A has a right inverse, then it is unique, since the previous equations have unique solutions. Reversed, if A is quadratic and $\rho(A) = n$, then it has a unique right inverse. Since the rank of A is n, the rank of its transpose is also n. Transposing the equation $YA = I$ we get $(YA)^T = A^TY^T = I^T = I$, and this system of equations has also a unique solution. Then $Y = X$ the unique inverse and it is denoted by A^{-1}.

If $AX = AY$, then it does not follows that $X = Y$. We have the example, that the product of two matrices non-zero is the zero matrix, and multiplying e.g. the first one from right by

the zero matrix results in the zero matrix as well. When does it follows from $AX = AY$ that $X = Y$? Since the columns both of AX and AY are linear combinations of the columns of A with constant multipiers from X and Y, these constant multipliers are unique only if the columns of A are linearly independent. Similarly, if $XB = YB$ then $X = Y$ follows from it only in the case when the rows of B are independent. Consequently, multiplying a matrix X from left by a matrix with independent columns or from right by a matrix with independent rows does not change the rank of X.
Obviously, if A is an invertible quadratic matrix, then multiplying the equation $AX = AY$ by the inverse of A from left we get $X = Y$. Consequently, multiplying a matrix by an invertible matrix from left or from right does not change the rank of the matrix.

4. Theorem. - $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^{-1})^T = (A^T)^{-1}$
- $(A^{-1})^* = (A^*)^{-1}$
- Invertible diagonal matrix has diagonal inverse.
- Invertible upper triangle matrix has upper triangle inverse.

11. Definition. If a quadratic matrix has a 1 in every column and in every row, and zeros everywhere else, it is called **permutation matrix**.

The rank of a permutation matrix is n, since it consists of the unit vectors.

Multiplying a matrix by a permutation matrix from left, the order of the rows will change in the way as the columns of the permutation matrix are changed from the unit matrix. Similarly, multiplying a matrix by a permutation matrix from right, the order of the colums will change in the way as the rows of the permutation matrix are changed from the unit matrix.

Let create a matrix A which differs from the unit matrix by only one element. Assume that this element is $a_{i,j} = \alpha \neq 0$. Multiplying a matrix B by A from left, the j-th row of B is added to the i-th row. This A is invertible, and its inverse is a matrix with $-\alpha$ on the place of $a_{i,j}$.

Corollaries:
- A system of linear equations can be written in matrix form: $A\mathbf{x} = \mathbf{b}$.
- The steps of transforming a coefficient matrix into row echelon form or reduced row echelon form is equivalent by multiplying this coefficient matrix by invertible matrices described above. Now, we have a second explanation for the fact that the Gauss-elimination steps are equivalent transformations.
- If the coefficient matrix is quadratic and it has an inverse, then the solution can be written in the form: $\mathbf{x} = A^{-1}\mathbf{b}$.

There are several different types of matrices of special forms or special properties. Some of them:
-If $P = P \cdot P = P^2$, the so called **projector** or **idempotent** matrices;
-If $A = A^T$, the so called symmetric matrices;
- If $A^* = A$, the so called **Hermitian** or **selfadjoint** matrices;
- If $N^k \neq 0$, $(1 \leq k)$ but $N^{k+1} = 0$, then N is called **nilpotent**.

Obviously the 0 matrix has several special property, e.g. it is diagonal, upper triangle, lower triangle, symmetric, self adjoint, etc.

The real symmetric matrices are selfadjoint matrices as well.

If $A_1 = BB^T$, $A_2 = B^T B$, $C_1 = BB^*$, $C_2 = B^* B$, then A_1 and A_2 are symmetric, C_1, C_2 are selfadjoint matrices.

We will discuss these special matrices later.

Changing coordinates in the case of changing the base

Question: If we know the coordinates of a vector in the old basis, and we know the coordinates of the vectors of the new basis (of course with respect to the old basis, since before the change everything is given there), then how can we tell the coordinates of that vector with respect to this new basis?

Consider a vector x. Let be the elements of the old basis a_1, a_2, ... ,a_n, and the elements of the new basis b_1, b_2, ... ,b_n. If the coordinates of the vector x in the old basis are $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and in the new basis they are $(\beta_1, \beta_2, \ldots, \beta_n)$, then

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n = \beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_n b_n$$

Let formulate this equality with coordinates with respect to the old basis. The coordinates of the element of the old fundamental system a_i are such that the i-th is 1,the others are zero. So, the equality is:

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \alpha_1 + \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \alpha_2 + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \alpha_n = \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} \beta_1 + \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{bmatrix} \beta_2 + \cdots + \begin{bmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{nn} \end{bmatrix} \beta_n$$

Written in matrix form:

$$\begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

The left hand side are the coordinates of x in the old basis. The columns of the matrix B are linearly independent, so B is an invertable matrix, and multiplying both side by its inverse, we get the coordinates of x in the new basis:

$$B^{-1} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

If we consider this equality for $x = a_1$, then the left hand side is

$$B^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and this is the first column of the inverse of B. I.e. the columns of B contain the coordinates of the new basis with respect to the old one, and the columns of B^{-1} contain the coordinates of the old basis with respect to the new one.