Rough Syllabus, Part 2.
Linear system of equations

The general form of a linear system of equations is the following:

$$\begin{array}{ccccccccc} a_{1,1}\ x_1 & + & a_{1,2}\ x_2 & + & \dots & + & a_{1,n}\ x_n & = & b_1 \\ a_{2,1}\ x_1 & + & a_{2,2}\ x_2 & + & \dots & + & a_{2,n}\ x_n & = & b_2 \\ \vdots & & & & & & \vdots & \vdots & \vdots \\ a_{m,1}\ x_1 & + & a_{m,2}\ x_2 & + & \dots & + & a_{m,n}\ x_n & = & b_m \end{array}$$

Here $a_{i,j}$ are the coefficients, the unknowns are x_j. The first index (row index) of $a_{i,j}$ shows from which equation is it, and the second index (column index) shows that which unknown has this coefficient.

A transformation of this system is called **identical transformation** if the original and the transformed system have exactly the same solutions.
Obviously, exchanging the order of the equations, the multiplication of an equation by a nonzero number and adding a constant multiple of an equation to an other one are identical transformations.

Gauss elimination method, row echelon form

Since in the following procedure we never change the order of the unknowns, only maybe the order of equations, we do not have to display the unknowns, it is enough to know their coefficients. So our table has the form

$$\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_2 \\ \vdots & & & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_m \end{array}$$

The left hand part of this table created from the coefficients of the unknowns is called the **coefficient matrix** of the system of equations. We will discuss ”matrices” later, but it is easiest to refer to the coefficients in this way.

Exchanging two rows of the table is exchanging the order of equations. Since this is an identical transformation, it is ”useless” to remember to the original order of equations, so we keep the indexes belonging to the positions of the coefficients in the table.

Let assume, that there is an equation in which the coefficient of x_1 is not zero, and put this one into the first place. The coefficient $a_{1,1}$ is called the **pivot** element of this step of calculations. This equation can be multiplied by $\frac{1}{a_{1,1}}$, so the coefficient of x_1 will be 1. Then we add an appropriate constant multiple of this equation to the others so, that the coefficients of x_1 become 0. Then in the second and in the later equations we have only the unknowns x_2, x_3,... x_n. If in these equations there is one in which the coefficient of x_2 is not zero, then we put this one on the place of the second one, and we repeat the previous procedure with these remaining equations. If in these equations all coefficients of x_2 are zero, it means that in the remaining system we do have x_2. Then we continue the procedure with x_3 in the same way, and we continue with the further unknowns until it is possible, until we can choose nonzero pivot element. Assume, we could repeat this step k times. Then in the table of coefficients we have k rows in each the first nonzero element is 1, and below it every coefficient is zero. This form of the system is called **row echelon form**, and the 1's in the first nonzero positions are the **leading 1's**.

E.g. Assume that we have five equation with 7 unknowns, and our procedure finished with the following table:

$$\begin{array}{ccccccc|c}
1 & c_{1,2} & c_{1,3} & c_{1,4} & c_{1,5} & c_{1,6} & c_{1,7} & d_1 \\
0 & 1 & c_{2,3} & c_{2,4} & c_{2,5} & c_{2,6} & c_{2,7} & d_2 \\
0 & 0 & 0 & 1 & c_{3,5} & c_{3,6} & c_{3,7} & d_3 \\
0 & 0 & 0 & 0 & 1 & c_{4,6} & c_{4,7} & d_4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & d_5
\end{array}$$

It means that performing identical transformations we get the following system of equations:

$$\begin{array}{ccccccccc}
x_1 & +c_{1,2}\ x_2 & +c_{1,3}\ x_3 & +c_{1,4}\ x_4 & +c_{1,5}\ x_5 & +c_{1,6}\ x_6 & +c_{1,7}\ x_7 & = & d_1 \\
 & x_2 & +c_{2,3}\ x_3 & +c_{2,4}\ x_4 & +c_{2,5}\ x_5 & +c_{2,6}\ x_6 & +c_{2,7}\ x_7 & = & d_2 \\
 & & & x_4 & +c_{3,5}\ x_5 & +c_{3,6}\ x_6 & +c_{3,7}\ x_7 & = & d_3 \\
 & & & & x_5 & +c_{4,6}\ x_6 & +c_{4,7}\ x_7 & = & d_4 \\
 & & & & & & 0 & = & d_5
\end{array}$$

Obviously, if $d_5 \neq 0$, then the system does not have any solution, since in the last equation every unknown has coefficient 0. If $d_5 = 0$, then this system has infinitely many solutions, since x_3, x_6, x_7 can have arbitrary values, from the fourth equation we can express uniquely x_5, then from the third one we can express x_4, etc. In this system variables x_3, x_6, x_7 are the so called **independent variables**, since their values can be chosen freely, x_1, x_2, x_4, x_5 are the **dependent variables**, since their values depend uniquely from the values of the independent variables.

If in the previous procedure we eliminate the nonzero element not only below the pivot element, but also above it, then we get the so called **reduced row echelon form**.

Considering the previous example the reduced row echelon form will be the following (now, coefficients $c_{i,j}$ and d_i are obviously different the ones given in the previous table):

$$\begin{array}{ccccccc|c}
1 & 0 & c_{1,3} & 0 & 0 & c_{1,6} & c_{1,7} & d_1 \\
0 & 1 & c_{2,3} & 0 & 0 & c_{2,6} & c_{2,7} & d_2 \\
0 & 0 & 0 & 1 & 0 & c_{3,6} & c_{3,7} & d_3 \\
0 & 0 & 0 & 0 & 1 & c_{4,6} & c_{4,7} & d_4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & d_5
\end{array}$$

The corresponding system of equations which is still equivalent with the original one is:

$$\begin{array}{cccccccc}
x_1 & & +c_{1,3}\ x_3 & & & +c_{1,6}\ x_6 & +c_{1,7}\ x_7 & = & d_1 \\
 & x_2 & +c_{2,3}\ x_3 & & & +c_{2,6}\ x_6 & +c_{2,7}\ x_7 & = & d_2 \\
 & & & x_4 & & +c_{3,6}\ x_6 & +c_{3,7}\ x_7 & = & d_3 \\
 & & & & x_5 & +c_{4,6}\ x_6 & +c_{4,7}\ x_7 & = & d_4 \\
 & & & & & & 0 & = & d_5
\end{array}$$

Here every dependent variable can be expressed directly with the independent ones.

Observe that if the columns of the table of coefficients are considered as $m-tuples$ written in column form, then the system can be written in the form:

$$\begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix} x_1 + \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

In vector form:

$$\mathbf{a_1}x_1 + \mathbf{a_2}x_2 + \cdots + \mathbf{a_n}x_n = \mathbf{b}$$

In previous exercises we have seen that $m - tuples$ with number elements form an m dimensional vector space over the field from which they have their elements, and they can be written in row, or in column form.

So, to solve a linear system of equations is equivalent to look for a linear combination of the column vectors of the coefficient matrix in order to express the column vector on the right hand side.

Since the reduced row echelon form is equivalent to the original system, written it in the same vector form we can see the following facts:
-If during the elimination method we could choose pivot elements k times, then there are k linearly independent vectors among the columns of the coefficient matrix, and they are the columns associated to the dependent variables.
-The columns of the independent variables can be expressed as a linear combination of the columns associated to the dependent variables, and the coefficients are as we can see from the table.
-If the vector created from the right hand side of the system is linearly independent from the columns of the coefficient matrix, then the system does not have any solution.
-If the system can be solved, and the columns of the coefficient matrix are not linearly independent, then the system has infinitely many solutions.

Considering our previous example we can observe, that the previous statements are obvious.

$$\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix} x_1 + \begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix} x_2 + \begin{bmatrix}c_{1,3}\\c_{2,3}\\0\\0\\0\end{bmatrix} x_3 + \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix} x_4 + \begin{bmatrix}0\\0\\0\\1\\0\end{bmatrix} x_5 + \begin{bmatrix}c_{1,6}\\c_{2,6}\\c_{3,6}\\c_{4,6}\\0\end{bmatrix} x_6 + \begin{bmatrix}c_{1,7}\\c_{2,7}\\c_{3,7}\\c_{4,7}\\0\end{bmatrix} x_7 = \begin{bmatrix}d_1\\d_2\\d_3\\d_4\\d_5\end{bmatrix}$$

Further observations about the coefficient matrix

Let consider the rows of the coefficient matrix (the rows of the table of coefficients). These are $n-tuples$, so can be considered as a system of vectors from an n-dimensional space. During the process to achive the reduced row echelon form we only add constant multiples of these vectors to each other. So, the following statements are valid:
-The rank of the this row vector system does not change during this procedure, and from the row echelon form we can see that it is equal to the number of the leading 1-s. -The subspace generated by the row vectors of the original coefficients is exactly the same as the subspace generated by the row vectors of the row echelon form, or by the reduced row echelon form.

Structure of the solution set of a linear system of equations

Structure of the solution set of a homogeneous linear system of equations

1. Definition. If on the right hand side every element is zero, then the system is called a **homogeneous** linear system of equations.

1. Theorem. A homogeneous linear system always has a solution, namely the identically zero solution.

From one side it is obvious, from the other side it follows from the fact, that the zero vector (right hand side column) is linearly dependent from any other vectors.

2. Definition. The identically zero solution of a homogeneous system is the so called **trivial** solution.

2. Theorem. A homogeneous linear system has non-trivial solution only if the rank of the column vector system of the coefficient matrix is less than the number of unknowns.

The number of the columns is equal to the number of unknowns. The rank is the maximal number of linearly independent vectors among them. If they are independent, then only the trivial linear combination can result in the zero vector.

The solutions have n elements, they can be considered as $n-tuples$.

3. Theorem. The solutions of a homogeneous linear system of equations form a linear subspace in the n dimensional space, and the dimension of this subspace is the number of the independent variables, i.e. the dimension of this subspace is $n-k$, if k is the rank of the column vectors of the coefficient matrix.

Proof:
The sum of solution vectors is also a solution vector:

$$\begin{array}{ccccccccc}
a_{1,1}\ (x_1+y_1) & + & a_{1,2}\ (x_2+y_2) & + & \dots & + & a_{1,n}\ (x_n+y_n) & = & 0 \\
a_{2,1}\ (x_1+y_1) & + & a_{2,2}\ (x_2+y_2) & + & \dots & + & a_{2,n}\ (x_n+y_n) & = & 0 \\
\vdots & & & & & & \vdots & \vdots & \vdots \\
a_{m,1}\ (x_1+y_1) & + & a_{m,2}\ (x_2+y_2) & + & \dots & + & a_{m,n}\ (x_n+y_n) & = & 0
\end{array}$$

with distributive multiplication and commutative addition

$$\begin{array}{lll}
(a_{1,1}\ x_1+a_{1,2}\ x_2+\dots+a_{1,n}\ x_n) & +(a_{1,1}\ y_1+a_{1,2}\ y_2+\dots+y_{1,n}\ y_n) & =0 \\
(a_{2,1}\ x_1+a_{2,2}\ x_2+\dots+a_{2,n}\ x_n) & +(a_{2,1}\ y_1+a_{2,2}\ y_2+\dots+a_{2,n}\ y_n) & =0 \\
\vdots & \vdots & \vdots \\
(a_{m,1}\ x_1+a_{m,2}\ x_2+\dots+a_{m,n}\ x_n) & +(a_{m,1}\ y_1+a_{m,2}\ y_2+\dots+a_{m,n}\ y_n) & =0
\end{array}$$

Constant multiple of a solution is also a solution:

$$
\begin{array}{lll}
(a_{1,1}\ \lambda x_1 + a_{1,2}\ \lambda x_2 + \cdots + a_{1,n}\ \lambda x_n) & = \lambda(a_{1,1}\ x_1 + a_{1,2}\ x_2 + \cdots + y_{1,n}\ x_n) & = 0 \\
(a_{2,1}\ \lambda x_1 + a_{2,2}\ \lambda x_2 + \cdots + a_{2,n}\ \lambda x_n) & = \lambda(a_{2,1}\ x_1 + a_{2,2}\ x_2 + \cdots + a_{2,n}\ x_n) & = 0 \\
\vdots & \vdots & \vdots \\
(a_{m,1}\ \lambda x_1 + a_{m,2}\ \lambda x_2 + \cdots + a_{m,n}\ \lambda x_n) & = (a_{m,1}\ x_1 + a_{m,2}\ x_2 + \cdots + a_{m,n}\ x_n) & = 0
\end{array}
$$

The statement of the theorem about the rank of this subspace is obvious if we observe the structure of the solutions what we get from the reduced row echelon form. Instead of applying general notations, let see our previous example.

Let denote the freely chosen values of the independent variables by $x_3 = r$, $x_6 = s$, $x_7 = t$. The dependent ones x_1, x_2, x_4 and x_5 can be expressed by them. So, the solution is:

$$
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} -c_{1,3} \\ -c_{2,3} \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} r + \begin{bmatrix} -c_{1,6} \\ -c_{2,6} \\ 0 \\ -c_{3,6} \\ -c_{4,6} \\ 1 \\ 0 \end{bmatrix} s + \begin{bmatrix} -c_{1,7} \\ -c_{2,7} \\ 0 \\ -c_{3,7} \\ -c_{4,7} \\ 0 \\ 1 \end{bmatrix} t
$$

We have the linear combination of so many linearly independent vectors, how many independent variables exist, i.e. all possible linear combinations of these vectors. They are independent, since each of these vectors has a 1 on the position of one of the independent variables, and 0 on the positions of the others. These vectors form a basis of the subspace of the solutions.

3. Definition. The solution set of a homogeneous system of linear equations given in a form as the linear combination of linearly independent vectors is called the **general solution** of the equation.

We will give an other explanation to the dimension of the solution set of a homogeneous linear system of equations after dicussing the scalar product in linear spaces.

Structure of the solution set of an inhomogeneous linear system of equations

4. Definition. If on the right hand side not every element is zero, then the system is called a **inhomogeneous** linear system of equations.

4. Theorem. An inhomogeneous system of linear equations has a solution only if the column of the right hand side is a linear combination of the columns of the coefficient matrix.

Proof: It is an obvious consequence of the vector form of the system.

5. Definition. If the inhomogeneous system can be solved, then a given solution is called a **particular solutionn of the system.**

5. Theorem. Two solutions of an inhomogeneous system differ from each other by a solution of the corresponding homogeneous system.

Proof: The solutions are $n-tuples$. Assume, $(x_1, x_2, \dots x_n)$ and $(y_1, y_2, \dots y_n)$ are solutions. Then substituting their difference in the left side of the system we get

$$\begin{array}{ccccccc}
a_{1,1}\ (x_1-y_1) & + & a_{1,2}\ (x_2-y_2) & + & \dots & + & a_{1,n}\ (x_n-y_n) \\
a_{2,1}\ (x_1-y_1) & + & a_{2,2}\ (x_2-y_2) & + & \dots & + & a_{2,n}\ (x_n-y_n) \\
\vdots & & \vdots & & & & \vdots \\
a_{m,1}\ (x_1-y_1) & + & a_{m,2}\ (x_2-y_2) & + & \dots & + & a_{m,n}\ (x_n-y_n)
\end{array}$$

with distributive multiplication and commutative addition

$$\begin{array}{rcll}
(a_{1,1}\ x_1 + a_{1,2}\ x_2 + \dots + a_{1,n}\ x_n) & - & (a_{1,1}\ y_1 + a_{1,2}\ y_2 + \dots + y_{1,n}\ y_n) & = b_1 - b_1 = 0 \\
(a_{2,1}\ x_1 + a_{2,2}\ x_2 + \dots + a_{2,n}\ x_n) & - & (a_{2,1}\ y_1 + a_{2,2}\ y_2 + \dots + a_{2,n}\ y_n) & = b_2 - b_2 = 0 \\
 & \vdots & & \quad\vdots \\
(a_{m,1}\ x_1 + a_{m,2}\ x_2 + \dots + a_{m,n}\ x_n) & - & (a_{m,1}\ y_1 + a_{m,2}\ y_2 + \dots + a_{m,n}\ y_n) & = b_m - b_m = 0
\end{array}$$

Corollary:

6. Theorem. If an inhomogeneous system of equations can be solved, then the general solution of it is the sum of a particular solution and the general solution of the corresponding homogeneous system.

Corollary:

7. Theorem. An inhomogeneous system of equations has a unique solution if it can be solved, and the corresponding homogeneous system has only the trivial solution, i.e. the columns of the coefficient matrix are linearly independent vectors.