1 Systems Of Linear Equations and Matrices

1.1 Systems Of Linear Equations

In this section you'll learn what Systems Of Linear Equations are and how to solve them.

Remember that equations of the form

$$a_1x + a_2y = b, \quad \text{for } a_1, a_2 \in \mathbb{R}\backslash\{0\}, b \in \mathbb{R}$$

describe lines in a 2-dimensional $(x - y)$ coordinate system. This is called a linear equation in x and y.
More generally

Definition 1
A *linear equation in n variables*, $x_1, x_2, \ldots, x_n$, is given by

$$a_1x_1 + a_2x_2 + \ldots a_nx_n = b$$

where $a_1, a_2, \ldots, a_n, b \in \mathbb{R}$.

Example 1
$3x_1 + 2x_2 = 9$, $3x_1 - 2x_2 + x_3 = 25$ are linear.
But $3x_1^2 + 2x_2 = 9$ and $3x_1 - 2x_2 + \sqrt{x_3} = 25$ are not linear.
They have x_1^2, and $\sqrt{x_3}$ included which are non linear terms in the variables.

Definition 2

1. A set of linear equations in the variables $x_1, x_2, \ldots, x_n$ is called a *system of linear equations* or a *linear system*.
2. A sequence of numbers $s_1, s_2, \ldots s_n$ is called a *solution* of the system if $x_1 = s1, x_2 = s_2, \ldots, x_n = s_n$ is a solution of every equation in the system.

Example 2

1. Consider the following linear system

$$\begin{aligned}&3x_1 + 2x_2 - 3x_3 = 10\\&x_1 - x_2 + x_3 = 2\end{aligned}$$

Then $x_1 = 3, x_2 = 2, x_3 = 1$ is a solution of the linear system, because

$$\begin{aligned}&3(3) + 2(2) - 3(1) = 10\\&3 - 2 + 1 = 2\end{aligned}$$

However $x_1 = 1, x_2 = 1, x_3 = 1$ is not a solution, because $3(1) + 2(1) - 3(1) = 2 \neq 10$.

2. Not all linear systems have solutions

   For example

   $$\begin{aligned} 3x_1 + 3x_2 &= 9 \\ 3x_1 + 3x_2 &= 3 \end{aligned}$$

   does not have a solution, the 2 equations contradict each other.

3. Other linear systems have infinite many solutions

   $$\begin{aligned} 3x_1 + 3x_2 &= 9 \\ x_1 + x_2 &= 3 \end{aligned}$$

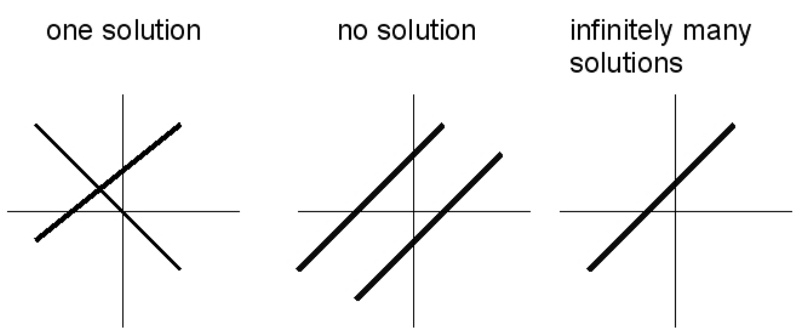
   For all $t \in \mathbb{R}$ $x_1 = t$ and $x_2 = 3 - t$ are solutions of the linear system.

   Check:

   $$\begin{aligned} 3t + 3(3-t) &= 3t + 9 - 3t = 9 \\ t + (3-t) &= t + 3 - t = 3 \end{aligned}$$

   To find just one solution, e.g. choose $x_1 = t = 3, x_2 = 3 - t = 3 - 3 = 0$, then this is one solution of the linear system.

Geometrically in 2 dimensions:



We will later prove that for every linear system exactly one of three possibilities hold

1. has no solution
2. has exactly one solution, or
3. has infinitely many solutions

Definition 3
A linear system that has no solution is called inconsistent
A linear system with at least one solution is called consistent.

In order to abbreviate the writing (mathematicians are lazy, when it comes to writing) matrices are used to represent linear systems.

Example 3
The following linear system

$$\begin{aligned} 3x_1 + 2x_2 - 3x_3 &= 10 \\ x_1 - x_2 + x_3 &= 2 \\ 4x_1 + 2x_2 &= 16 \end{aligned}$$

can be represented, by just listing the constants in the system, but the location has to be kept in mind. The augmented matrix representing this linear system is

$$\begin{bmatrix} 3 & 2 & -3 & 10 \\ 1 & -1 & 1 & 2 \\ 4 & 2 & 0 & 16 \end{bmatrix}$$

In general:
An arbitrary system of m linear equations in n unknowns can be written as

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

where $x_1, x_2, \dots, x_n$ are the unknowns, and the a's and b's denote the constants.
The indexes on the coefficients are specifying the location in the system. The first index indicates the equation, and the second index gives the unknown the coefficient multiplies.
a_{53} is the multiplier for variable x_3 in equation 5.

Definition 4

1. A *matrix* is a rectangular array of numbers.

2. The *augmented matrix* for the linear system (6) is

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix} \tag{2}$$

The first task is to learn, how to solve linear equations.
You will learn an algorithm to do so. This is very essential to be mastered to succeed in this course. We will use the algorithm over and over for solving different types of problems.

Systems of linear equations are solved by manipulating the system through operations, that do not change the solution, but make finding the solution easier.

Elementary row operations:

1. Multiply an equation by a constant (not zero).
2. Interchange two equations.
3. Add a multiple of one equation to another.

Use the operations to eliminate unknowns from equations to make solving the system easier.
To save some writing we could do the same operations with the augmented matrix

Example 4

The linear system and its augmented matrix

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ 3x_1 + 2x_2 - 3x_3 &= 10 \\ 4x_1 + 2x_2 &= 16 \end{aligned} \qquad \begin{bmatrix} 1 & -1 & 1 & 2 \\ 3 & 2 & -3 & 10 \\ 4 & 2 & 0 & 16 \end{bmatrix}$$

1. Add -3 times the first equation to the second

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ 5x_2 - 6x_3 &= 4 \\ 4x_1 + 2x_2 &= 16 \end{aligned}$$

1. -3 (1st row) + (2nd row)

$$\begin{bmatrix} 1 & -1 & 1 & 2 \\ 0 & 5 & -6 & 14 \\ 4 & 2 & 0 & 16 \end{bmatrix}$$

2. Add -4 times the first equation to the third

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ 5x_2 - 6x_3 &= 4 \\ 6x_2 - 4x_3 &= 8 \end{aligned}$$

2. -4 (1st row) + (3rd row)

$$\begin{bmatrix} 1 & -1 & 1 & 2 \\ 0 & 5 & -6 & 4 \\ 0 & 6 & -4 & 8 \end{bmatrix}$$

3. Multiply the second equation by 1/5

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ x_2 - 6/5x_3 &= 4/5 \\ 6x_2 - 4x_3 &= 8 \end{aligned}$$

3. 1/5 (2nd row)

$$\begin{bmatrix} 1 & -1 & 1 & 2 \\ 0 & 1 & -6/5 & 4/5 \\ 0 & 6 & -4 & 8 \end{bmatrix}$$

4. Add -6 times the second equation to the third

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ x_2 - 6/5x_3 &= 4/5 \\ 16/5x_3 &= 16/5 \end{aligned}$$

4. -6 (2nd row) +(3rd row)

$$\begin{bmatrix} 1 & -1 & 1 & 2 \\ 0 & 1 & -6/5 & 4/5 \\ 0 & 0 & 16/5 & 16/5 \end{bmatrix}$$

5. Multiply the third equation by 5/16

$$\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ x_2 - 6/5x_3 &= 4/5 \\ x_3 &= 1 \end{aligned}$$

5. 5/16 (3rd row)

$$\begin{bmatrix} 1 & -1 & 1 & 2 \\ 0 & 1 & -6/5 & 4/5 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

6. Add -1 times the third equation to the first

$$\begin{aligned} x_1 - x_2 &= 1 \\ x_2 - 6/5x_3 &= 4/5 \\ x_3 &= 1 \end{aligned}$$

6. -(3rd row)+ (1st row)

$$\begin{bmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & -6/5 & 4/5 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

7. Add -6/5 times the third equation to the second

$x_1 - x_2 = 1$
$x_2 = 10/5$
$x_3 = 1$

7. -6/5(3rd row)+ (2nd row)

$$\begin{bmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

8. Add 1 times the second equation to the first

$x_1 = 3$
$x_2 = 2$
$x_3 = 1$

8. (2nd row)+ (1st row)

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

Now it is obvious that the solution is $x_1 = 3, x_2 = 2, x_3 = 1$. This system has exactly one solution.

Example 5

1. Assume you have one equation in 3 variables

$$3x + 3y - 3z = 9$$

What are the solutions of this equation? Since we have only one equation, but three variables, we can choose freely (3-1) variables. Let $s, t \in \mathbb{R}$, then $y = s, z = t, 3x + 3s - 3t = 9 \Leftrightarrow y = s, z = t, x = (9 - 3s + 3t)/3 = 3 - s + t$ is a solution for the equation.

E.g. choose $y = 1, z = 1$, then with $x = 3 - 1 + 1 =$ is one possible solution.

2.

$$\begin{aligned} x_1 - 2x_2 &= 5 \\ -2x_1 + 4x_2 &= -10 \end{aligned}$$

Use the augmented matrix to solve this linear system

$$\begin{bmatrix} 1 & -2 & 5 \\ -2 & 4 & -10 \end{bmatrix} \rightarrow_{2(1st\ row)+(2nd\ row)} \begin{bmatrix} 1 & -2 & 5 \\ 0 & 0 & 0 \end{bmatrix}$$

The last row translates to $0 = 0$, which is always true, and gives no restriction on the unknowns solving the system. This line can be ignored when determining the solution.

(2 variables - 1 equation) For $s \in \mathbb{R}$, $x_2 = s, x_1 = 5 + 2s$ is a solution of the linear system above. The system has infinite many solutions.

3.

$$\begin{aligned} x_1 - 2x_2 &= 5 \\ -2x_1 + 4x_2 &= 5 \end{aligned}$$

Use the augmented matrix to solve this linear system

$$\begin{bmatrix} 1 & -2 & 5 \\ -2 & 4 & 5 \end{bmatrix} \rightarrow_{2(1st\ row)+(2nd\ row)} \begin{bmatrix} 1 & -2 & 5 \\ 0 & 0 & 15 \end{bmatrix}$$

The last row translates to $0 = 15$, which is not true. No choice of the variables will lead to a solution of the linear system.

This system has no solution.

1.2 Gaussian Elimination

Definition 5

1. A matrix is in row-echelon form, if and only if

   (a) The leading coefficient of each nonzero row is one. (leading 1)

   (b) All rows consisting entirely of zeros are grouped at the bottom of the matrix.

   (c) The leading nonzero coefficient of a row is always strictly to the right of the leading coefficient of the row above it.

2. A matrix is in reduced row-echelon form, if and only if

   the matrix is in row-echelon form, and

   (d) each column that contains a leading 1 has zero everywhere else in that column.

Theorem 1
Every matrix in reduced row-echelon form is also in row-echelon form.

Proof:
Assume A is any matrix in reduced row-echelon form.
Therefore properties (a)-(d) from Definition 5 hold for A.
Therefore particularly properties (a)-(c) hold for A.
Therefore according to the definition A is in row-echelon form.
End of proof.
(Use therefore = "$\Rightarrow$")

Gaussian Elimination is an algorithm to transform a matrix into a equivalent matrix in row-echelon form.
Gauss-Jordan elimination is an algorithm to transform a matrix into a equivalent matrix in reduced row-echelon form.

Example 6

1.

$$\begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -10 \end{bmatrix}$$

Gives solution $x_1 = 5, x_2 = 2, x_3 = -10$

2.

$$\begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 2 & 2 \end{bmatrix}$$

x_1 and x_2 correspond to leading 1s in the matrix, they are called *pivots* or *leading variables*.
The nonleading variables are called *free*, here x_3.
Free variables can be assigned arbitrary value from $\mathbb{R}$, say s. Then the leading variables can be solved in s.
This gives solutions $x_1 = 5, x_2 = 2 - 2s, x_3 = s$, for any $s \in \mathbb{R}$.

1.
$$\left[\begin{array}{cccc} 1 & 0 & 0 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & -10 \end{array}\right]$$

The last equation translates to $0x_3 = -10$, so $0 = -10$. This can not be satisfied, the linear system has no solution.

Gaussian Elimination

step 1 Locate the leftmost column that is not entirely 0.

step 2 Exchange the top row with another row, if necessary, to bring a non zero entry to the top of the column from step 1.

step 3 If the entry is not one, say it is a, then multiply the row by $1/a$.

step 4 Add suitable multiples of the top row to the rows bellow so that all entries below the leading one become zeros.

step 5 Now start over with step 1 for the matrix excluding row 1.

The matrix is now in row echelon form

Gauss-Jordan Elimination
Step 1 to Step 5, and

step 6 Beginning with the last nonzero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading 1s.

After step 6 the matrix is in reduced row-echelon form.

Example 7

Consider a linear system with the following augmented matrix

$$\left[\begin{array}{cccc} 0 & -1 & -2 & -16 \\ 1 & 2 & 1 & 8 \\ 2 & 2 & 0 & 2 \end{array}\right]$$

step 1 First column

step 2
$$\left[\begin{array}{cccc} 1 & 2 & 1 & 8 \\ 0 & -1 & -2 & -16 \\ 2 & 2 & 0 & 2 \end{array}\right]$$

step 3 the entry is 1

step 4 -2(row 1)+(row 3)

$$\begin{bmatrix} 1 & 2 & 1 & 8 \\ 0 & -1 & -2 & -16 \\ 0 & -2 & -2 & -14 \end{bmatrix}$$

step 5 start over with step 1.

step 1 second column

step 2 it is not zero

step 3 -1(row 2)

$$\begin{bmatrix} 1 & 2 & 1 & 8 \\ 0 & 1 & 2 & 16 \\ 0 & -2 & -2 & -14 \end{bmatrix}$$

step 4 2(row 2)+ (row 3)

$$\begin{bmatrix} 1 & 2 & 1 & 8 \\ 0 & 1 & 2 & 16 \\ 0 & 0 & 2 & 18 \end{bmatrix}$$

step 5 start over with step 1

step 1 third column

step 2 1/2(row 3)

$$\begin{bmatrix} 1 & 2 & 1 & 8 \\ 0 & 1 & 2 & 16 \\ 0 & 0 & 1 & 9 \end{bmatrix}$$

This ends the Gaussian elimination, the matrix is in row-echelon form.
Using back substitution we can use the linear system from here.
The third row translates to $x_3 = 9$.
The second row translates to $x_2 + 2(x_3) = 16$, using $x_3 = 9$, we get
$x_2 = 16 - 2(9) = -2$
The first row translates to $x_1 + 2x_2 + x_3 = 8$, with $x_2 = -2, x_3 = 9$, we get
$x_1 = 8 - 2(-2) - 9 = 3$.
The only solution is $x_1 = 3, x_2 = -2, x_3 = 9$.

Instead of using back substitution we could have found the reduced row-echelon form by using step 6.

step 6 -2(row 3) + row 2, -1(row3)+(row1)

$$\begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 9 \end{bmatrix}$$

-2(row 2)+ (row 1)

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 9 \end{bmatrix}$$

This is the reduced row-echelon form, which translate immediately into the solution $x_1 = 3, x_2 = -2, x_3 = 9$.

Definition 6
A linear system is said to be *homogeneous* if in

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}$$

$b_1 = b_2 = \ldots = b_m = 0$

Theorem 2
All homogeneous linear systems have at least the *trivial* solution $x_1 = x_2 = \ldots = x_n = 0$.

Proof:
Substitute $x_1 = x_2 = \ldots = x_n = 0$ into the first equation, and find

$$a_{11}0 + a_{12}0 + \ldots + a_{1n}0 = 0$$

Equation 1 is satisfied. In the same way we see that all the other equations will be satisfied, and $x_1 = x_2 = \ldots = x_n = 0$ is a solution of the homogeneous system.

Definition 7
All solutions different from $x_1 = x_2 = \ldots = x_n = 0$ are called nontrivial.

Theorem 3
Every homogeneous linear system with more variables than equations has infinite many solutions.

Proof:(not presented in class)
$m < n$ Then the reduced row echelon form of the augmented matrix for the linear system looks like

$$\begin{bmatrix} 1 & & & & & & S_1 & 0 \\ & 1 & & & & & S_2 & 0 \\ & & 1 & & & & S_3 & 0 \\ & & & \ddots & & & \vdots & \vdots \\ & & & & 1 & & S_{m-1} & 0 \\ & & & & & 1 & S_m & 0 \end{bmatrix}$$

Where S_j are the coefficients for the $n - m > 0$ free variables in row j, $1 \leqslant j \leqslant m$.

Most important to observe, there have to be free variables, which can be set to arbitrary numbers in $\mathbb{R}$, and each choice results in a different solution, so that there are must be infinite many different solutions.

Solutions of homogeneous linear systems, can also be found using Gaussian or Gauss-Jordan Elimination

Example 8

The augmented matrix of a homogeneous linear system is

$$\begin{bmatrix} 1 & 2 & -3 & 0 \\ 2 & -2 & 0 & 0 \\ 3 & -4 & 1 & 0 \end{bmatrix} \rightarrow_{-2(r1)+r2,-3(r1)+r3} \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & -6 & 6 & 0 \\ 0 & -8 & 8 & 0 \end{bmatrix}$$

$$\rightarrow_{-1/6(r2)} \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -8 & 8 & 0 \end{bmatrix} \rightarrow_{8(r2)+r3} \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

(backward substitution) The third row only yields $0 = 0$, from the second row we get $x_2 = x_3$, from the first row, we then get $x_1 + 2x_3 - 3x_3 = 0 \Leftrightarrow x_1 = x_3$.
For any $s \in \mathbb{R}$ let $x_3 = s$ (free variable), then $x_1 = x_2 = x_3 = s$ is a solution of the linear system. With $s = 0$, we get the trivial solution.

1.3 Matrices and Matrix Operations

In Definition 4, we gave the definition of a matrix to be a rectangular array of numbers.

Definition 8

1. The numbers in the matrix are called the *entries* in the matrix.
2. A matrix with m rows and n columns is called an $m \times n$ matrix.
3. A matrix with only one column ($m \times 1$ matrix) is called a column matrix or column vector.
4. A matrix with only one row ($1 \times n$ matrix) is called a row matrix or row vector.

Denotation

- Capital letters are used to denote matrices, and lower case letters are used to indicate numerical quantities (scalars).
- The entry in row i and column j of a matrix A is denoted by a_{ij} or $(A)_{ij}$.
- A general $m \times n$ matrix can be written as

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \tag{3}$$

- When compactness of notation is desired, the preceding matrix can be written as $[a_{ij}]_{m\times n}$ or $[a_{ij}]$.
- Denote vectors by lowercase boldface letters $\mathbf{a}$ being a $1 \times n$ (row) vector, and $\mathbf{b}$ being a $m \times 1$ column vector can be written as

$$\mathbf{a} = [a_1, a_2, \dots, a_n], \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Example 9

- $A = \begin{bmatrix} 1 & 2 & -3 & 0 \\ 2 & -2 & 0 & 0 \\ 3 & -4 & 1 & 0 \end{bmatrix}$ is a 3×4 matrix. $a_{23} = (A)_{23} = -2$ and $a_{32} = (A)_{32} = -4$
- $\mathbf{a} = [2, 4, -3]$ is a 1×3 vector, with $a_2 = 4$.

Definition 9
An $n \times n$ matrix A is called a square matrix of order n.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \tag{4}$$

The entries $a_{11}, a_{22}, \dots a_{nn}$ define the main diagonal of A.

Definition 10
If $A = [a_{ij}]$ and $B = [b_{ij}]$ have the same size, then $A = B \Leftrightarrow a_{ij} = b_{ij}$ for all i and j.

This definition says, that for two matrices to be equal they have to be of the same size, AND all entries have to be equal.

Example 10
Let

$$A = \begin{bmatrix} -3 & 1 \\ 5 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} -3 & 1 \\ s & 2 \end{bmatrix}, \quad C = \begin{bmatrix} -3 & 1 \\ 4 & t \end{bmatrix}$$

Then $A = B \Leftrightarrow s = 5$,
$A \neq C$, because $a_{21} \neq c_{21}$
$B = C \Leftrightarrow s = 4, t = 2$

Definition 11
If $A = [a_{ij}]$ and $B = [b_{ij}]$ have the same size, then the sum of A and B, $A + B$, is defined by

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij} = a_{ij} + b_{ij} \quad \forall ij$$

and the difference of A and B, $A - B$, is defined through

$$(A - B)_{ij} = (A)_{ij} - (B)_{ij} = a_{ij} - b_{ij} \quad \forall ij$$

This definition says, that matrices of the same size are added and subtracted by adding and subtracting the corresponding entries, respectively.

Example 11
Let

$$A = \begin{bmatrix} -3 & 1 \\ 5 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} -3 & 1 \\ 4 & t \\ 5 & 2 \end{bmatrix}$$

Then

$$A + B = \begin{bmatrix} -1 & 0 \\ 4 & 6 \end{bmatrix}, \quad A - B = \begin{bmatrix} -5 & 2 \\ 6 & -2 \end{bmatrix}$$

C can neither be added to nor subtracted from A and B because the sizes are not equal.

Theorem 4
Let A, B, C be matrices of the same size.

(a) $A + B = B + A$ (commutative law for addition)

(b) $A + (B + C) = (A + B) + C$ (associative law for addition)

Proof:

(a) Since the size of the sum of two matrices is equal to the size of the matrices being added, the matrices on the right and left are of the same size.

We will now show that all entries of the two matrices are the same, then according to Definition 10, the matrices are equal.

For all choices of i, j

$$\begin{aligned} (A+B)_{ij} &\overset{Def.11}{=} (A)_{ij} + (B)_{ij} \\ &= (B)_{ij} + (A)_{ij} \\ &\overset{Def.11}{=} (B+A)_{ij} \end{aligned}$$

(b) DIY

Definition 12
If $A = [a_{ij}]$ and $c \in \mathbb{R}$, then the product of A and c, cA, is defined by

$$(cA)_{ij} = c(A)_{ij} = ca_{ij} \quad \forall ij$$

Example 12
Let $c = -1$ and

$$A = \begin{bmatrix} -3 & 1 \\ 5 & 2 \end{bmatrix}$$

then

$$cA = \begin{bmatrix} 3 & -1 \\ -5 & -2 \end{bmatrix}$$

Theorem 5
Let $a, b \in \mathbb{R}$, and B, C matrices of the same size, then

(a) $(a + b)C = aC + bC$

(b) $(a - b)C = aC - bC$

(c) $a(bC) = (ab)C$

(d) $a(B + C) = aB + aC$

(e) $a(B - C) = aB - aC$

Proof:

(a) Since multiplication with a scalar and adding two matrices has no effect on the size of the matrix, the matrices on the right and left have the same size.

We will next show that all entries of the two matrices are the same, then according to Definition 10, the matrices are equal. For any i, j

$$\begin{aligned} ((a+b)C)_{ij} &\overset{Def.12}{=} (a+b)(C)_{ij} \\ &= a(C)_{ij} + b(C)_{ij} \\ &\overset{Def.12}{=} (aC)_{ij} + (bC)_i j \\ &\overset{Def.11}{=} (aC+bC)_i j \end{aligned}$$

(c) Since multiplication with a scalar has no effect on the size of a matrix, the matrices on the right and left have the same size. Now show that all entries are the same. For any i, j

$$\begin{aligned} (a(bC))_{ij} &\overset{Def.12}{=} a(bC)_{ij} \\ &\overset{Def.12}{=} (ab)(C)_{ij} \\ &\overset{Def.12}{=} ((ab)C)_{ij} \end{aligned}$$

- Try the other proofs.

Definition 13
Let $A_1, A_2, \ldots, A_n$ be matrices of the same size, and $c_1, c_2, \ldots, c_n \in \mathbb{R}$, then an expression of the form

$$c_1 A_1 + c_2 A_2 + \ldots + c_n A_n$$

is called a linear combination in $A_1, A_2, \ldots, A_n$ with coefficients $c_1, c_2, \ldots, c_n$.

Example 13
Let

$$A = \begin{bmatrix} -3 & 1 \\ 5 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} -4 & 4 \\ -2 & 2 \end{bmatrix}$$

then

$$2A - 3B + \frac{1}{2}C = \begin{bmatrix} -6 & 2 \\ 10 & 4 \end{bmatrix} + \begin{bmatrix} -6 & 3 \\ 3 & -12 \end{bmatrix} + \begin{bmatrix} -2 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -14 & 7 \\ 12 & -7 \end{bmatrix}$$

is a linear combination in A, B, C with coefficients $2, -3, 1/2$.

Definition 14
If $A = [a_{ij}]$ is an $m \times r$ matrix and $B = [b_{ij}]$ is an $r \times n$ matrix, then the product of A and B, AB, is a $m \times n$ matrix and, is defined through

$$(AB)_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj} \quad \forall i, j$$

Example 14

- Let

$$A = \begin{bmatrix} -3 & 1 & 2 \\ 5 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 & 3 \\ -1 & 4 & -1 \\ 3 & -2 & 1 \end{bmatrix}$$

  A is a 2×3 matrix and B is a 3×3 matrix. Since the number of columns of A= number of rows of B, the product is defined and a 2×3 matrix

$$AB = \begin{bmatrix} -1 & 3 & -8 \\ 11 & -1 & 14 \end{bmatrix}$$

  1st row of A, 1st column of B: $(AB)_{11} = -3(2) + 1(-1) + 2(3) = -1$
  2nd row of A, 1st column of B: $(AB)_{21} = 5(2) + 2(-1) + 1(3) = 11$

  1st row of A, 2nd column of B: $(AB)_{12} = -3(-1) + 1(4) + 2(-2) = 3$
  2nd row of A, 2nd column of B: $(AB)_{22} = 5(-1) + 2(4) + 1(-2) = -1$

  1st row of A, 3rd column of B: $(AB)_{13} = -3(3) + 1(-1) + 2(1) = -8$
  2nd row of A, 3rd column of B: $(AB)_{23} = 5(3) + 2(-1) + 1(1) = 14$

- One special case is the multiplication of a row vector $\mathbf{a}_{1\times n}$ with a column vector $\mathbf{b}_{n\times 1}$, by definition $\mathbf{ab}$ is a 1×1 matrix, i.e. a scalar.

$$\mathbf{ab} = \sum_{i=1}^{n} a_i b_i$$

  This will be later defined to be the dot-product of two vectors.

- We can partition matrices and conceive them as a an array of row or column vectors. If A is a $m \times r$ and B is a $r \times n$ matrix, then

$$A = \left[\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1r} \\ \hline a_{21} & a_{22} & \dots & a_{2r} \\ \hline \vdots & \vdots & & \vdots \\ \hline a_{m1} & a_{m2} & \dots & a_{mr} \end{array}\right] = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}, \quad B = \left[\begin{array}{c|c|c|c} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{r1} & b_{r2} & \dots & b_{rn} \end{array}\right] = \left[\begin{array}{c|c|c|c} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_n \end{array}\right]$$

  Then the ij entry of the product of A and B is the product of row vector i with column vector j:
  $(AB)_{ij} = \mathbf{a}_i \mathbf{b}_j$

  Or one could partition the product matrix and observe, that

$$AB = \left[\begin{array}{c|c|c|c} A\mathbf{b}_1 & A\mathbf{b}_2 & \dots & A\mathbf{b}_n \end{array}\right] = \begin{bmatrix} \mathbf{a}_1 B \\ \mathbf{a}_2 B \\ \vdots \\ \mathbf{a}_m B \end{bmatrix}$$

- If $A = [\mathbf{a_1}|\mathbf{a_2}|\cdots|\mathbf{a_n}$ is a $m \times n$ matrix and $\mathbf{x}$ is a column vector of length n, then $A\mathbf{x} = x_1\mathbf{a_1} + x_2\mathbf{a_2} + \ldots + x_n\mathbf{a_n}$ is a linear combination of the vectors making up A, or

$$A\mathbf{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix}$$

  Therefore a linear system in variables $x_1, \ldots, x_n$ can be written in matrix form as $A\mathbf{x} = \mathbf{b}$, where A is the coefficient matrix.

Theorem 6
Assuming that the sizes of the matrices A, B, C are such that the indicated operations are defined and $a \in \mathbb{R}$, then

(a) $A(BC) = (AB)C$ (associative law for multiplication)

(b) $A(B + C) = AB + AC$ (left distribution law)

(c) $(A + B)C = AC + BC$ (right distribution law)

(d) $A(B - C) = AB - AC$

(e) $(A - B)C = AC - BC$

(f) $a(BC) = (aB)C = B(aC)$

Proof: I will only show how to proof one of the properties.
(b)$A(B + C) = AB + AC$
To show this property, we need to prove according to Definition 10 that the matrices have the same size and all corresponding entries for the two matrices on the right and left are the same.
In order to add two matrices, they have to have the same size. Let B, C be $r \times n$ matrices, therefore $B+C$ according to Definition 11 is a $r \times n$ matrix. For the products $A(B+C), AB, AC$ to be defined assume A is a $m \times r$ matrix, then $A(B + C), AB, AC$ are all $m \times n$ matrices, therefore $AB + AC$ is a $m \times n$ matrix, therefore the sizes of the two matrices are equal.
We found that the matrices on the right and the left have the same size. Now we will prove that the corresponding entries are equal.
Choose any ij then

$$\begin{aligned}
(A(B+C))_{ij} &\overset{Def.14}{=} \sum_{k=1}^{r} (A)_{ik}(B+C)_{kj} \\
&\overset{Def.11}{=} \sum_{k=1}^{r} (A)_{ik}((B)_{kj} + (C)_{kj}) \\
&= \sum_{k=1}^{r} ((A)_{ik}(B)_{kj} + (A)_{ik}(C)_{kj}) \qquad \text{Distributive law for scalars} \\
&= \sum_{k=1}^{r} ((A)_{ik}(B)_{kj} + \sum_{k=1}^{r} (A)_{ik}(C)_{kj}) \quad \text{Algebra}
\end{aligned}$$

$$\overset{Def.14}{=} (AB)_{ij} + (AC)_{ij}$$

$$\overset{Def.11}{=} (AB + AC)_{ij}$$

This proves that every entry in matrix $A(B+C)$ equals the corresponding entry in matrix $AB+AC$, which finalizes the proof.

Definition 15
If A is a $m \times n$ matrix, then the transpose of A, A^T, is a $n \times m$ matrix with $(A^T)_{ij} = (A)_{ij}$

Transposing a matrix results in a matrix where rows and columns are exchanged.

Example 15

- $$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \text{then } A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

- $$A = \begin{bmatrix} -3 & 1 & 2 \end{bmatrix} \text{ then } A^T = \begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix}$$

Theorem 7
If the sizes of the matrices are such that the stated operations can be performed then

(a) $(A^T)^T = A$

(b) $(A + B)^T = A^T + B^T$ and $(A - B)^T = A^T - B^T$

(c) For $k \in \mathbb{R}$: $(kA)^T = kA^T$

(d) $(AB)^T = B^T A^T$

Proof:

(a) Let A be a square matrix of size n, then A^T is a square matrix of size n, therefore $(A^T)^T$ is a square matrix of size n. Therefore the matrices on the left and right have the same size. Now we will prove that all entries are the same. For any i, j

$$((A^T)^T)_{ij} = (A^T)_{ji} = (A)_{ij}$$

Therefore the entries are all the same and the equation holds.

(b) (first part) Let A, B square matrices of size n, then $A + B$ is a square matrix of size n, therefore $A^T, B^T, (A+B)^T$ are square matrices of size n, and also $A^T + B^T$ is a square matrix of size n. The size of the matrices on the right and left are equal, prove now that all entries are equal. For any i, j

$$((A+B)^T)_{ij} \overset{Def.15}{=} (A+B)_{ji} \overset{Def.11}{=} (A)_{ji} + (B)_{ji} \overset{Def.15}{=} (A^T)_{ij} + (B^T)_{ij} \overset{Def.11}{=} (A^T + B^T)_{ij}$$

Therefore the entries are all the same and the equation holds.

Definition 16
If A is a square matrix of size n, then the trace of A, $tr(a)$, is defined by

$$tr(a) = a_{11} + a_{22} + \ldots + a_{nn} = \sum_{i=1}^{n} a_{ii}$$

The trace of a square matrix is the total of the entries on the main diagonal

Example 16

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -3 & 4 & -2 \end{bmatrix}, \text{ then } \ tr(A) = 1 + 5 + (-2) = 4$$

theorem trace?

1.4 Inverses

Definition 17
The zero matrix, 0, of size $m \times n$ is the matrix with $(0)_{ij} = 0$ for all index combinations ij.

Theorem 8
Assuming that the sizes of the matrices are such that the indicated operations are defined, the following equalities hold

(a) $A + 0 = 0 + A = A$

(b) $A - A = 0$

(c) $0 - A = -A$

(d) $A0 = 0$, $0A = 0$

Proof: only for part (d) first part Let A be a $m \times r$ matrix and $O_{r\times n}$ be a $r \times n$ matrix then according to Definition 14 $A0$ is a $m \times n$ matrix. Now show that the entries of this matrix are all zero.
For any ij

$$(A0_{r\times n})_{ij} = \sum_{k=1}^{r} a_{ik} 0_{kj} = \sum_{k=1}^{r} a_i k 0 = \sum_{k=1}^{r} 0 = 0$$

All entries of $A0$ are 0, therefore $A0 = 0$.

Definition 18
The identity matrix of size n, I_n, is a $n \times n$ matrix with all entries 0, beside the entries on the main diagonal being 1.

Example 17

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 & 0 \\ 0 & \ldots & 0 & 0 & 1 \end{bmatrix}$$

Theorem 9
For all matrices A of size $m \times n$

(a) $AI_n = A$

(b) $I_m A = A$

Proof: only part (a)
If A is a $m \times n$ matrix then AI_n is defined and a $m \times n$ matrix, therefore the sizes of the matrices on the right and left are equal.

Now prove that the entries are all the same. For any ij

$$(AI_n)_{ij} = \sum_{k=1}^{n} a_{ik}(I_n)kj = \sum_{k \neq j} a_{ik}0 + a_{ij}1 = a_{ij} = (A)_{ij}$$

Since the size of the matrices are the same and the entries are all equal, the matrices are equal, therefore the equation holds.

Theorem 10
If R is the reduced row echelon form of an $n \times n$ matrix A, then either R has a row of zeros or $R = I_n$.

Proof: (text book pg 42) (not in class)

Definition 19
If A is a square matrix, and if a matrix B of the same size exists such that

$$AB = BA = I_n$$

then A is said to be invertible and B is called an inverse of A.
If no such matrix exists, then A is said to be singular.

Example 18

- $$A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \quad B = \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}$$

  B is the inverse of A. To show that this is true we need to confirm the properties of the inverse matrix.

  $$AB = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

  and

  $$BA = \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

- The matrix

  $$A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$$

  is singular (i.e. has no inverse). Let B be any matrix of size 2×2

  $$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

  Then

  $$BA = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} b_{11} + 2b_{21} & 0 \\ b_{21} + 2b_{22} & 0 \end{bmatrix} \neq I_2$$

  because the entry in row 2 column 2 is always 0, which should be 1 for the matrix to be the identity.

Theorem 11
If matrices B and C are inverses of A, then $B = C$. (The inverse of an invertible matrix is unique).

Proof:
Assume that B and C are inverses of A, then

$$\begin{array}{rrcll} & BA & = & I & \text{multiply both sides with } C \text{ from the right} \\ & (BA)C & = & C & \text{associativity of multiplication} \\ \Leftrightarrow & B(AC) & = & C & AC = I (C \text{ is an inverse of } A) \\ \Leftrightarrow & BI & = & C & \\ \Leftrightarrow & B & = & C & \end{array}$$

From the assumptions we conclude that $B = C$, which concludes the proof.

Denotation
Because the inverse is unique, we will denote the inverse of matrix A by A^{-1}.

Theorem 12
If A and B are invertible matrices of the same size, then AB is invertible and

$$(AB)^{-1} = B^{-1}A^{-1}$$

Proof:
Let A and B be invertible matrices of the same size, then

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = I$$

and

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I$$

We found that the matrix $B^{-1}A^{-1}$ has all properties of an inverse. Since the inverse is unique, it must be the inverse and AB must be invertible.

Theorem 13
A product of a finite number of invertible matrices is invertible, and the inverse of the product is the product of he inverses in reverse order.

Proof by induction not here.

Theorem 14
The matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is invertible if and only if $(ad - bc) \neq 0$, and in this case the inverse of A is given by

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Proof: DIY

Definition 20
If A is a square matrix, then for $n \in \mathbb{N}$

$$A^0 = I \quad \text{and } A^n = \underbrace{AA \cdots A}_{n \text{ factors}}$$

and if A is invertible

$$A^{-n} = \underbrace{A^{-1}A^{-1} \cdots A^{-1}}_{n \text{ factors}}$$

Theorem 15
If A is a square matrix and $r, s \in \mathbb{N}$, then

$$A^r A^s = A^{r+s} \text{and} (A^r)^s = A^{rs}$$

Theorem 16
If A is an invertible matrix, then

(a) A^{-1} is invertible and $(A^{-1})^{-1} = A$

(b) For all $n \in \mathbb{N}$, A^n is invertible and $(A^n)^{-1} = (A^{-1})^n$

(c) For all $k \in \mathbb{R}, k \neq 0$, kA is invertible and $(kA)^{-1} = 1/kA^{-1}$

Proof: only part (a), try (b) and (c) by yourself.
Since $A^{-1}A = I_n$ and $AA^{-1} = I_n$, we found a matrix $B(= A)$, so that $A^{-1}B = I$ and $BA^{-1} = I$, then according to Definition 19 A^{-1} is invertible and A is the inverse.

Theorem 17
If A is an invertible matrix, then A^T is invertible and

$$(A^T)^{-1} = (A^{-1})^T$$

Proof:
Observe that $A^T(A^{-1})^T = A^{-1}A = I)$ and $(A^{-1})^T A^T = AA^{-1} = I)$, i.e. we found a matrix of the same size as A^T with $A^T B = BA^T = I$, so A^T is invertible with $(A^T)^{-1} = A{-}1^T$.