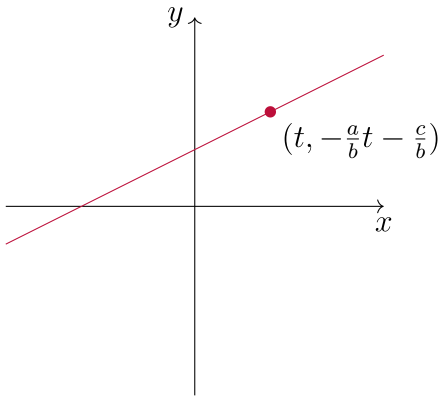
Chapter 1

Linear equations

A linear equation in one variable is on the form $ax + b = 0$ with a, b known values and x unknown. Such an equation has a unique solution $x = -b/a$. In two variables it is of the form $ax + by + c = 0$ with an infinite amount of solutions, namely the line

$$y = -\frac{a}{b}x - c$$

which can be visualized as a line in the xy-plane.



That is, for every number t we have a solution of the equation given by

$$x = t, \quad y = -\frac{a}{b}t - \frac{c}{b}.$$

In these cases we often call x a **free variable**. We could also have chosen y as a free variable, and then the solutions would have been on the form

$$x = -\frac{b}{a}r - \frac{c}{a}, \quad y = r$$

for every number r.

A linear equation in three variables is of the form $ax + by + cz + d = 0$ and the solutions will be points on a plane. Here we have two free variables, which we can choose to be y and z. Then the solutions will be on the form

$$x = -\frac{b}{a}t - \frac{c}{a}r - \frac{d}{a}, \quad y = t, \quad z = r$$

for every pair of numbers t and r.

Definition 1.0.1

A **linear equation** is of the form

$$a_1x_1 + a_2x_2 + \cdots + a_{n-1}x_{n-1} + a_nx_n = b$$

where $a_1, a_2, \ldots, a_{n-1}, a_n$ are known values, and $x_1, x_2, \ldots, x_{n-1}, x_n$ are unknown. The numbers a_i, for $i = 1, 2, \ldots, n$, are called the **coefficients** of the equation.

When you stumble upon linear equations in the wild you will more often than not have several equations describing the same unknowns x_i. That is, you have more than one constraint. This brings us to the concept of systems of linear equations.

Definition 1.0.2

A system of m linear equations in n variables is on the form of a collection of m linear equations with variables $x_1, x_2, \ldots, x_n$.

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

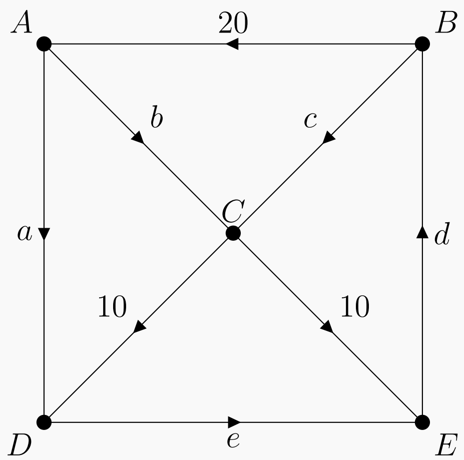
Example 1.0.1

When working with some quantity flowing through a network, say electric current, water or money, we can be interested in how much of the given quantity that flows through a specific part of the network.

A network consist of **junctions** or **nodes** with **branches** connecting them. Each branch has a given direction of flow and the underlying assumption is that the sum of everything flowing into a node should equal the amount flowing out.

This assumption is equal to Kirchoff's current law, when the network is a curcuit and the quantity is an electric current.

Let us look at the following network with some unknown quantities



The unknown quantities here are a, b, c, d. In order to find these, we use the assumption that what flows into the nodes A, B, C, D, E should equal what flows out of them to set up a system of linear equations.

$$\begin{aligned}
A: &\quad a + b = 20 \\
B: &\quad d - c = 20 \\
C: &\quad b + c = 20 \\
D: &\quad e - a = 10 \\
E: &\quad d - e = 10
\end{aligned}$$

Let us hold of on solving the system for a moment.

Example 1.0.2

Find the line $y = ax + b$ that intersect the points $(5, 8/2)$ and $(2, 5/2)$.

By substituting the values in the line equation, we obtain the following system of two linear equations with a and b unknown.

$$\begin{cases} 5a & + & b & = & \frac{8}{2} \\ 2a & + & b & = & \frac{5}{2} \end{cases}$$

Now we can either use substitution to solve this set, or we can add and subtract the equations. Let us try both methods.

Substitution: From the first equation we have that

$$b = \frac{8}{2} - 5a$$

substituting this value with b in the second equation gives us

$$2a + \frac{8}{2} - 5a = \frac{5}{2}$$

which in turn tells us that $a = \frac{1}{2}$. Now, by substituting this in our expression for b gives us $b = \frac{3}{2}$. That is the line is given by

$$y = \frac{1}{2}x + \frac{3}{2}$$

Adding and Subtracting equations: If we replace the second equation with the equation we get by subtracting the original first equation from the original second equation, we get

$$\begin{cases} 5a & + & b & = & \frac{8}{2} \\ -3a & + & 0 & = & -\frac{3}{2} \end{cases}$$

Now we can divide the second equation by -3:

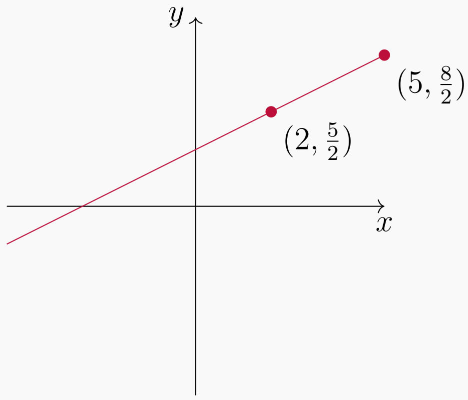
$$\begin{cases} 5a & + & b & = & \frac{8}{2} \\ a & + & 0 & = & \frac{1}{2} \end{cases}$$

before replacing the first equation, with the equation we get by subtracting five times the second equation from the first

$$\begin{cases} 0 & + & b & = & \frac{3}{2} \\ a & + & 0 & = & \frac{1}{2} \end{cases}$$

Reading out from this resulting system we get that the line is given by

$$y = \frac{1}{2} + \frac{3}{2}$$



1.1 Finding solutions

The last method used above is the one we are mostly interested in. It uses the same ideas as the method called Gaussian elimination, which is partly described through three elementary operations on the equations that do not change the solutions.

Remark 1.1.1

The three following elementary operations on systems do not change the resulting solutions.

1. Interchange any pair of equations
2. Multiply any equation by a nonzero number
3. Replace any equation by its sum with a multiple of any other equation.

Augmented matrix

When working with systems of linear equations there is a lot of symbols that can confuse us and obfuscate what we are doing. Hence we would like to simplify them and work only with the essential information.

Definition 1.1.1

The **coefficient matrix** of a system of linear equation

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

is the array of numbers

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

consisting of m rows and n columns, where the columns contain every coefficient of the equations, such that

- column one has the coefficients of x_1,
  column two has the coefficient of x_2,
  ...
- row one has the coefficients from the first linear equation,
  row two has the coefficients from the second linear equation,
  ...

Let us also introduce the following **column vectors**

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

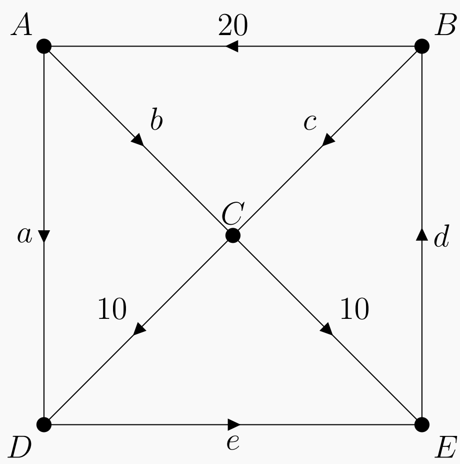
and define $A\mathbf{x}$ to be equal to

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

We can see that if A is the coefficient matrix of a system of linear equations and $\mathbf{b}$ is the column vector defined above, then the **matrix equation** $A\mathbf{x} = \mathbf{b}$ encodes the same information as the system.

Example 1.1.1

Remember that the network



gave us the following system of linear equations

$$\begin{aligned} A: &\quad a+b=20 \\ B: &\quad d-c=20 \\ C: &\quad b+c=20 \\ D: &\quad e-a=10 \\ E: &\quad d-e=10 \end{aligned}$$

this in turn gives us

$$\overbrace{\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}}^{A} \overbrace{\begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix}}^{\mathbf{x}} = \overbrace{\begin{bmatrix} 20 \\ 20 \\ 20 \\ 10 \\ 10 \end{bmatrix}}^{\mathbf{b}}$$

Remark 1.1.2

If we have the matrix equation

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

and translate it to a linear equation $ax + by = 0$, we might recognize that the left hand side is nothing but the dot product of the two-dimensional vectors

$$\begin{bmatrix} a \\ b \end{bmatrix} \text{ and } \begin{bmatrix} x \\ y \end{bmatrix}.$$

Correspondingly for

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

we see that the left hand side is equal to $ax + by + cz$ which is the dot product of the three-dimensional vectors

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \text{ and } \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

In fact, we have that the **dot-product of n-dimensional vectors**

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

is defined to be exactly

$$\mathbf{x} \cdot \mathbf{y} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Hence, we can see $A\mathbf{x}$ as being given as

$$\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{x} \\ \mathbf{a}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{a}_m \cdot \mathbf{x} \end{bmatrix}$$

where $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are the rows of A seen as vectors.

When we are doing elementary operations on systems of linear equations, the only thing that changes in the corresponding matrix equation $A\mathbf{x} = \mathbf{b}$ are the rows in A and rows in $\mathbf{b}$. The changes in the matrix equation is equivalent to those on the system, and we can therefore work on the matrix equation instead of the linear system. Also, since $\mathbf{x}$ do not change, we can omit this in the calculation.

Definition 1.1.2

The augmented matrix of the system of linear equations

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

is given by

$$\left[\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array}\right]$$

The vertical line tells us what is on the right and left of the equality signs in the corresponding system. Now, the elementary operations can be formulated as

1. Interchange any pair of rows
2. Multiply any row by a nonzero number
3. Replace any row by its sum with a multiple of any other row.

Let us highlight this connection by going back to the equations in example 1.0.2.

Example 1.1.2

We were solving the system

$$\begin{cases} 5a & + & b & = & \frac{8}{2} \\ 2a & + & b & = & \frac{5}{2} \end{cases}$$

The augmented matrix of this system is

$$\left[\begin{array}{cc|c} 5 & 1 & \frac{8}{2} \\ 2 & 1 & \frac{5}{2} \end{array}\right]$$

$$\begin{cases} 5a & + & b & = & \frac{8}{2} \\ 2a & + & b & = & \frac{5}{2} \end{cases} \qquad \left[\begin{array}{cc|c} 5 & 1 & \frac{8}{2} \\ 2 & 1 & \frac{5}{2} \end{array}\right]$$

$$\xrightarrow{II=II-I} \begin{cases} 5a & + & b & = & \frac{8}{2} \\ -3a & + & 0 & = & -\frac{3}{2} \end{cases} \qquad \xrightarrow{II=II-I} \left[\begin{array}{cc|c} 5 & 1 & \frac{8}{2} \\ -3 & 0 & -\frac{3}{2} \end{array}\right]$$

$$\xrightarrow{II=\frac{1}{3}\cdot II} \begin{cases} 5a & + & b & = & \frac{8}{2} \\ a & + & 0 & = & \frac{1}{2} \end{cases} \qquad \xrightarrow{II=\frac{1}{3}\cdot II} \left[\begin{array}{cc|c} 5 & 1 & \frac{8}{2} \\ 1 & 0 & \frac{3}{2} \end{array}\right]$$

$$\xrightarrow{II=I-5\cdot II} \begin{cases} 0 & + & b & = & \frac{3}{2} \\ a & + & 0 & = & \frac{1}{2} \end{cases} \qquad \xrightarrow{I=I-5II} \left[\begin{array}{cc|c} 0 & 1 & \frac{3}{2} \\ 1 & 0 & \frac{1}{2} \end{array}\right]$$

Definition 1.1.3

A matrix is in **row echelon form** if the following is satisfied:

1. All nonzero rows are above any rows of all zeros.
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zeros.

If a matris in echelon form also satisfies the following conditions, then it is in **reduced row echelon form**:

4. Each leading entry in each nonzero row is 1.
5. Each leading 1 is the only nonzero entry in its column.

The leading entries in the row echelon forms are called **pivot elements** and the columns they appear in are called **pivot columns**.

The following matrix is in row echelon form

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

and the following matrix is in reduced row echelon form

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

If this had been the reduced row echelon form of the augmented matrix of a linear equation

$$\begin{array}{ccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & a_{1,3}x_3 & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & a_{2,3}x_3 & = & b_2 \\ a_{3,1}x_1 & + & a_{3,2}x_2 & + & a_{3,3}x_3 & = & b_3 \end{array}$$

then we can read out from the reduced row echelon form

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

that the solution is given by $x_1 = 0$, $x_2 = 1$ and $x_3 = 0$.

Solution strategy

Remark 1.1.3

The general solution strategy for solving systems of linear equations is to

1. Find the augmented matrix of the system.
2. Use the elementary row operations to find the reduced row echelon form.
3. Read out the solution.

One can also in step two choose to only find the row echelon form and use substitution to find the other solutions.

Exactly one solution: If all the columns on the left of the vertical line contain a pivot element in the reduced row echelon form, then you have a unique solution to the system.

Infinite solutions: If you have columns on the left of the vertical line that contain non-zero entries, but are not pivot columns in the reduced row echelon form, then they correspond to **free variables** and you will have an infinite amount of solutions.

No Solutions: If you in the course of reduction get a row with only zero entries on the left of the vertical line, but non-zero entries on the right, then the system is called inconsistent and there will be no solution. This correspond to the equation $0 = 1$ which is ludicrous.

In example 1.0.2 we got a unique solution to the system. Now, let us look at the network in example 1.0.1 and observe that here we obtain an infinite amount of solutions.

Example 1.1.3 Infinite solutions

In example 1.0.1 we found the following matrix equation

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} 20 \\ 20 \\ 20 \\ 10 \\ 10 \end{bmatrix}$$

we solve for (a, b, c, d, e) by reducing the following augmented matrix into reduced row echelon form

$$\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20 \\ 0 & 0 & -1 & 1 & 0 & 20 \\ 0 & 1 & 1 & 0 & 0 & 20 \\ -1 & 0 & 0 & 0 & 1 & 10 \\ 0 & 0 & 0 & 1 & -1 & 10 \end{array}\right] \xrightarrow{IV=IV+I} \left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20 \\ 0 & 0 & -1 & 1 & 0 & 20 \\ 0 & 1 & 1 & 0 & 0 & 20 \\ 0 & 1 & 0 & 0 & 1 & 30 \\ 0 & 0 & 0 & 1 & -1 & 10 \end{array}\right]$$

$$\xrightarrow{II\leftrightarrow III} \left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20 \\ 0 & 1 & 0 & 0 & 1 & 30 \\ 0 & 1 & 1 & 0 & 0 & 20 \\ 0 & 0 & -1 & 1 & 0 & 20 \\ 0 & 0 & 0 & 1 & -1 & 10 \end{array}\right] \xrightarrow{III=III-II} \left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20 \\ 0 & 1 & 0 & 0 & 1 & 30 \\ 0 & 0 & 1 & 0 & 1 & -10 \\ 0 & 0 & -1 & 1 & 0 & 20 \\ 0 & 0 & 0 & 1 & -1 & 10 \end{array}\right]$$

$$\xrightarrow{IV=IV+III} \left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20 \\ 0 & 1 & 0 & 0 & 1 & 30 \\ 0 & 0 & 1 & 0 & -1 & -10 \\ 0 & 0 & 0 & 1 & -1 & 10 \\ 0 & 0 & 0 & 1 & -1 & 10 \end{array}\right] \xrightarrow{V=V-VI} \left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 20 \\ 0 & 1 & 0 & 0 & 1 & 30 \\ 0 & 0 & 1 & 0 & -1 & -10 \\ 0 & 0 & 0 & 1 & -1 & 10 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{I=I-II} \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & -1 & -10 \\ 0 & 1 & 0 & 0 & 1 & 30 \\ 0 & 0 & 1 & 0 & -1 & -10 \\ 0 & 0 & 0 & 1 & -1 & 10 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

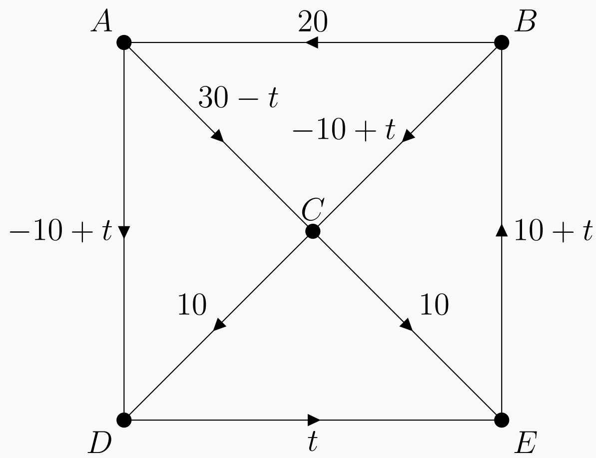
From the reduced form we read out that e is a free variable and the general solution of the system is

$$a = -10 + t, \ b = 30 - t, \ c = -10 + t, \ d = 10 + t, \ e = t$$

for any number t, or written in vector form

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} -10+t \\ 30-t \\ -10+t \\ 10+t \\ t \end{bmatrix} = \begin{bmatrix} -10 \\ 30 \\ -10 \\ 10 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

where t times a vector means that we multiply each entry with t and addition of vectors is component wise.



The minus sign in the network tells us that the quantity moves in opposite direction of what is indicated by the arrow.

Example 1.1.4 No solutions

Let us try to solve the following system

$$\begin{cases} \quad x - \;\; 2y = 1 \\ -5x + 10y = -1 \end{cases}$$

We reduce the augmented matrix

$$\left[\begin{array}{cc|c} 1 & -2 & 1 \\ -5 & 10 & -1 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -2 & 1 \\ 0 & 0 & 4 \end{array}\right]$$

The last matrix correspond to the system

$$\begin{cases} x + 2y = 1 \\ 0x + 0y = 4 \end{cases}$$

The equation $0x + 0y = 4$ is equal to $0 = 4$, which can't be true, no matter the value we give x og y. The system has no solutions.

Let us look at a few more examples to warm up our thinking cap.

Example 1.1.5

We wish to solve the system

$$\begin{array}{ccccccc} 2x_1 & + & 3x_2 & + & x_3 & = & 16 \\ 3x_1 & + & x_2 & + & 4x_3 & = & 27 \\ 4x_1 & + & x_2 & + & x_3 & = & 18. \end{array}$$

Re-writing this as an augmented matrix, we have

$$\left[\begin{array}{ccc|c} 2 & 3 & 1 & 16 \\ 3 & 1 & 4 & 27 \\ 4 & 1 & 1 & 18 \end{array}\right].$$

Performing the row operations,

$$\left[\begin{array}{ccc|c} 2 & 3 & 1 & 16 \\ 3 & 1 & 4 & 27 \\ 4 & 1 & 1 & 18 \end{array}\right] \xrightarrow[III=III-2\cdot I]{II=2\cdot II-3\cdot I} \left[\begin{array}{ccc|c} 2 & 3 & 1 & 16 \\ 0 & -7 & 5 & 6 \\ 0 & -5 & -1 & -14 \end{array}\right]$$

$$\xrightarrow{III=7\cdot III-5\cdot II} \left[\begin{array}{ccc|c} 2 & 3 & 1 & 16 \\ 0 & -7 & 5 & 6 \\ 0 & 0 & -32 & -128 \end{array}\right]$$

$$\xrightarrow{III=-\frac{1}{32}\cdot III} \left[\begin{array}{ccc|c} 2 & 3 & 1 & 16 \\ 0 & -7 & 5 & 6 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

$$\xrightarrow[II=II-5\cdot III]{I=I-III} \left[\begin{array}{ccc|c} 2 & 3 & 0 & 12 \\ 0 & -7 & 0 & -14 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

Rewriting this as a system, we have

$$\begin{array}{ccccccc} 2x_1 & + & 3x_2 & + & 0x_3 & = & 12 \\ 0x_1 & - & 7x_2 & + & 0x_3 & = & -14 \\ 0x_1 & + & 0x_2 & + & 1x_3 & = & 4. \end{array}$$

Thus $x_3 = 4$, $x_2 = 2$. Plugging these into the first equation, we have $x_1 = 3$. Written in vector form we have

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$$

Example 1.1.6

We are given the following system to solve:

$$\begin{cases} x_1 + 3x_2 + 2x_3 + 3x_4 = 16 \\ x_1 + 3x_2 + 3x_3 + x_4 = 21 \\ 2x_1 + 6x_2 + 4x_3 + 6x_4 = 32 \end{cases}$$

As usual we make the augmented matrix and start reducing,

$$\left[\begin{array}{cccc|c} 1 & 3 & 2 & 3 & 16 \\ 1 & 3 & 3 & 1 & 21 \\ 2 & 6 & 4 & 6 & 32 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 3 & 2 & 3 & 16 \\ 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$
$$\sim \left[\begin{array}{cccc|c} 1 & 3 & 0 & 7 & 6 \\ 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The last matrix correspond to the following system

$$\begin{cases} x_1 + 3x_2 + 7x_4 = 6 \\ \qquad\quad x_3 - 2x_4 = 5 \end{cases}$$

We have chosen to ignore the equation corresponding to the row of zeros. This is becuase it only tells us that $0x_1 + 0x_2 + 0x_3 + 0x_4 = 0$, or in other words $0 = 0$. This equation is satisfied regardless of the values for x_i, so it doesn't really tell us anything useful.

After moving everything except x_1 and x_3 to the right, we have:

$$\begin{cases} x_1 = -3x_2 - 7x_4 + 6 \\ x_3 = 2x_4 + 5 \end{cases}$$

That is, x_2 and x_4 are free variables. We can therefore find solutions of the system by setting x_2 and x_4 to be whatever we should please, and then using the equations to find what x_1 and x_3 needs to be.

For example, for $x_2 = 0$ and $x_4 = 1$, we get:

$$\begin{cases} x_1 = -3 \cdot 0 - 7 \cdot 1 + 6 = -1 \\ x_2 = 0 \\ x_3 = 2 \cdot 1 + 5 = 7 \\ x_4 = 1 \end{cases}$$

In order to describe all possible solutions, we set $x_2 = s$ and $x_4 = t$, for s and t arbitrary numbers. Then the general solution is:

$$\begin{cases} x_1 = -3s - 7t + 6 \\ x_2 = s \\ x_3 = 2t + 5 \\ x_4 = t \end{cases}$$

or written as vectors

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 5 \end{bmatrix} + s \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -7 \\ 0 \\ 2 \\ 1 \end{bmatrix}$$

Question 1.1.1

If we have a system of more variables than equation, and the system is solvable, will we always have free variables? Why?

Hint: We can at most have one pivot element in each row. Count rows and columns.

Some degree of freedom

When we are reducing a matrix, we have a certain degree of freedom. However, we can't do anything other than the three types of row operations. In what order and which rows we operate on are up to us though.

Example 1.1.7

We are reducing the following matrix:

$$\left[\begin{array}{cccc|c} 0 & 2 & 4 & 1 & 7 \\ 3 & 8 & 2 & 0 & 4 \\ 5 & 9 & 2 & 4 & 4 \end{array}\right]$$

Here we need to exchange the upper row with one of the other to get a pivot element in the right position. We can freely choose which row to move to the top.

We do also have some freedom when choosing the free variables. We can do as in example 1.1.6, that is to choose those variables which correspond to columns without pivot elements in the reduced form, but we can also choose other variables.

Example 1.1.8

In the example 1.1.6 we ended up with the two variables x_2 and x_4 being free. We could also have chosen x_1 and x_3 to be free.

We reduced the system to the following:

$$\begin{cases} x_1 = -3x_2 - 7x_4 + 6 \\ x_3 = 2x_4 + 5 \end{cases}$$

Solving the second equation for x_4:

$$x_4 = \frac{x_3 - 5}{2}$$

Substitution into the first equation gives:

$$\begin{aligned} x_2 &= \frac{-x_1 - 7x_4 + 6}{3} \\ &= \frac{-x_1 - \frac{7}{2}(x_3 - 5) + 6}{3} = -\frac{1}{3}x_1 - \frac{7}{6}x_3 + \frac{47}{6} \end{aligned}$$

Now, by setting $x_1 = s$ and $x_3 = t$, we obtain the following general solution:

$$\begin{cases} x_1 = s \\ x_2 = -\frac{1}{3}s - \frac{7}{6}t + \frac{47}{6} \\ x_3 = t \\ x_4 = \frac{1}{2}x_3 - \frac{5}{2} \end{cases}$$

Or in vector form:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{47}{6} \\ 0 \\ -\frac{5}{2} \end{bmatrix} + s \begin{bmatrix} 1 \\ -\frac{1}{3} \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -\frac{7}{6} \\ 1 \\ \frac{1}{2} \end{bmatrix}$$

At first glance this looks different from what we had in 1.1.6, but it describes the same solutions. (If you let $s = -1$ and $t = 7$, we have the same solution as when we had $x_2 = 0$ and $x_4 = 1$ in example 1.1.6).

⚠ Remark 1.1.4

Do note that even if we have an infinite amount of solution, every variable **can't** necessarily be chosen as free. For example, if a system of three variables is reduced to

$$\begin{aligned} x + y &= 3 \\ 2z &= 8 \end{aligned}$$

Then either x or y can be chosen as being free, but z can **not** be free.

Chapter 2

Vectors and matrices

2.1 Vectors

We now introduce the most common vector, the ones used to model the real world. An abstract notion of a vector needs to be based off the properties that these vectors satisfy.

Definition 2.1.1 Vector in $\mathbb{R}^n$

Let n be some natural number, i.e. positive integer. Denote by $\mathbb{R}^n = \mathbb{R}\times\mathbb{R}\times\cdots\mathbb{R}$ the set of all n-tuples of real numbers. We write these in a few common notations. For $\mathbf{v} \in \mathbb{R}^n$,

- $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ (**column vector notation**),
- $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ (**row vector notation**),
- $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ (**coordinate notation**).

Vectors are usually denoted by a bold lowercase letter, $\mathbf{v}$ in typed texts. If written by hand it is common to denote vectors by an arrow above, $\vec{v}$, or a line below, $\underline{v}$, instead of boldtype.

Remark 2.1.1

Unless explicitly told otherwise we will by a vector mean a column vector. Also, to save vertical space we will often write out column vectors as list of numbers. That is $\mathbf{v} = (1, 2, 3)$ will be the same as

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

We wish to explore some properties of vectors. Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. Then the following properties hold

- $\alpha \mathbf{v} = \alpha \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{bmatrix}$ (**scalar multiplication**),
- $\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$ (**vector addition**).

We also have some geometric properties of vectors. Note that vectors have a sense of direction and length,

- the **dot product/inner product** of two vectors is given by $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1 v_2 + u_2 v_2 + \ldots u_n v_n$,
- The **magnitude/length** of a vector is given by $|v| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} = \sqrt{\mathbf{v} \cdot \mathbf{v}}$,
- the angle between two vectors, θ is given by the relation $\mathbf{u} \cdot \mathbf{v} = |u||v| \cos(\theta)$,
- two vectors are orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$, or if the angle between them is a right angle.

A vector on the form

$$a\mathbf{x} + b\mathbf{y} = a \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} + b \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} ax_1 + by_1 \\ ax_2 + by_2 \\ \vdots \\ ax_m + by_m \end{bmatrix}$$

is a **linear combination** of the vectors $\mathbf{x}$ and $\mathbf{y}$. The numbers a and b is called scalars. If we have n vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$, then a vector on the form

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n$$

is a linear combination of the vectors $\mathbf{x}_i$ with scalars a_1.

If we have n vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$, we define the linear span, or

$$\operatorname{span}(\mathbf{x}_1, \dots, \mathbf{x}_n)$$

to be the set of all linear combinations of $\mathbf{x}_1, \dots, \mathbf{x}_n$, that is every possible scaled sum

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n$$

Example 2.1.1

The vector $\begin{bmatrix} 11 \\ 16 \\ 21 \end{bmatrix}$ is a linear combination of $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ og $\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$. since

$$3\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 11 \\ 16 \\ 21 \end{bmatrix}$$

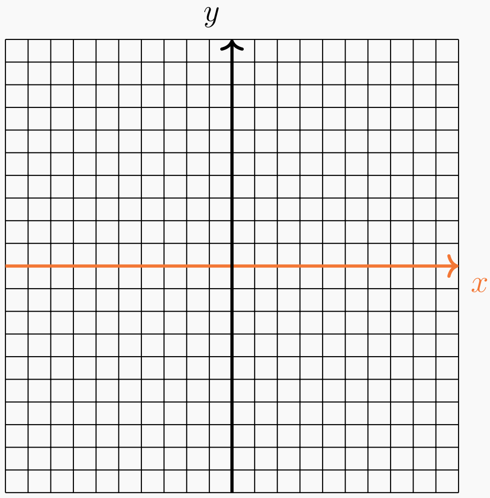
The span of $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ og $\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ are all vectors i $\mathbb{R}^3$ on the form

$$a\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + b\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix},$$

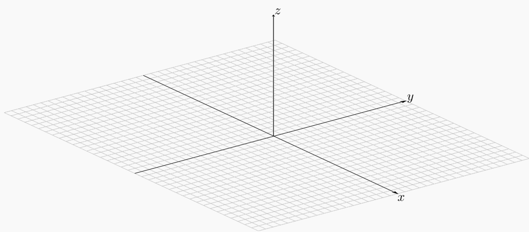
with a and b numbers.

In $\mathbb{R}^2$ and $\mathbb{R}^3$ such sets can be easily visualized.

Example 2.1.2



a) The span of the set $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$ is the x-axis in the xy-plane.



b) The span of the set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ in $\mathbb{R}^3$ (or the xyz-space) are all vectors on the form

$$\begin{bmatrix} a \\ b \\ 0 \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

for all values of a and b. This is the xy-plane.

Vector equations

Let us look back to the definitions leading up to augmented matrices, specifically the matrix equation $A\mathbf{x} = \mathbf{b}$. The left hand side of the equation was defined to be

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

and this can be split into a scaled sum of the column vectors of A

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

Remark 2.1.2

Observe that $A\mathbf{x}$ is the sum of the column vectors of A, scaled by the entries of $\mathbf{x}$, i.e.

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

Thus $A\mathbf{x} = \mathbf{b}$ can also be seen as a **vector equation**.

This gives us a new way to look at systems of linear equations: The problem is to find the scales x_i such that the linear combination of the matrix columns equals the right hand side.

Example 2.1.3

The vector equation

$$x \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + y \begin{bmatrix} 2 \\ 5 \\ 5 \end{bmatrix} + z \begin{bmatrix} -2 \\ 9 \\ -1 \end{bmatrix} = \begin{bmatrix} -5 \\ 33 \\ 0 \end{bmatrix}.$$

is equivalent to the system of equation

$$\begin{cases} x + 2y - 2z = -5 \\ x + 5y + 9z = 33 \\ 2x + 5y - z = 0 \end{cases}$$

and has a unique solution:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7 \\ -2 \\ 4 \end{bmatrix},$$

Verify yourselves that the following is in fact an equality

$$7\begin{bmatrix}1\\1\\2\end{bmatrix} - 2\begin{bmatrix}2\\5\\5\end{bmatrix} + 4\begin{bmatrix}-2\\9\\-1\end{bmatrix} = \begin{bmatrix}-5\\33\\0\end{bmatrix}.$$

Remark 2.1.3

If the right hand side of a matrix equation $A\mathbf{x} = \mathbf{b}$ is zero, i.e. $\mathbf{b} = \mathbf{0}$, then the equation is called **homogeneous**, and if it is nonzero, $\mathbf{b} \neq \mathbf{0}$, then the equation is called **non-homogeneous**.

The general solution of a matrix equation $A\mathbf{x} = \mathbf{b}$ is given by the sum of the general solution of the homogeneous case $\mathbf{x}_h$ and a particular solution of the non-homogeneous case $\mathbf{x}_p$,

$$\mathbf{x} = \mathbf{x}_h + \mathbf{x}_p.$$

Linear Transformations

When working with vectors we will often use a particular type of function on them that behaves nicely. Let us indulge in a small description of these.

Definition 2.1.2 Linear Transformation

We call a function $T : \mathbb{R}^n \to \mathbb{R}^m$ a **linear transformation** if it preserves scalar multiplication and vector addition, that is

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y),$$

for any $\alpha, \beta \in \mathbb{R}$ and $x, y \in \mathbb{R}^n$.

Example 2.1.4 Some linear transformations

The following are some examples of linear transformations:

- $\mathrm{Id} : \mathbb{R}^2 \to \mathbb{R}^2, \begin{bmatrix}x_1\\x_2\end{bmatrix} \mapsto \begin{bmatrix}x_1\\x_2\end{bmatrix}$ (**Identity Mapping**),
- $S : \mathbb{R}^2 \to \mathbb{R}^2, \begin{bmatrix}x_1\\x_2\end{bmatrix} \mapsto \begin{bmatrix}x_2\\x_1\end{bmatrix}$,
- $T : \mathbb{R}^3 \to \mathbb{R}^3, T(x_1, x_2, x_3) = (x_1, 2x_2, x_3)$,

- $P : \mathbb{R}^3 \to \mathbb{R}^2, P(x_1, x_2, x_3) = (x_1, x_2)$.

You can confirm these are linear transformations directly using the definition.

We also introduce an examples of a mapping that is **not** a linear transformation.

- $\tilde{T} : \mathbb{R}^2 \to \mathbb{R}^2, (x_1, x_2) \mapsto (x_1^2, x_2)$.

Linear transformations are particularly nice functions, however in the forms given they are not necessarily nice to work with. Is there a way we can write these in a nice way?

2.2 Matrices

We have for some time now worked with matrices without really going into the gritty details. Let us look a bit more closely on what they are. We will se that vectors are a special type of matrices, and that the operations possible on them extends nicely to matrices as well.

Definition 2.2.1

A **matrix** is a rectangular array of numbers. A matrix A having m rows and n columns is written as

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

The **size** of a matrix is the number of rows and columns, and we say that a matrix A is of size $m \times n$, or A is an $m \times n$-matrix, if it has m rows and n columns. Two matrices of equal size and with the same elements in corresponding positions are equal.

Remark 2.2.1

Another useful notion when working with matrices is to consider the matrix being built up of column vectors. If we let $\mathbf{a}_j$ be the column vector consisting of the elements of the jth column of a $m \times n$-matrix A, then we can write A as

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a_2} & \cdots & \mathbf{a_n} \end{bmatrix}$$

$$\begin{array}{c c}
 & \begin{array}{ccccc} \phantom{a_{1,1}} & & \text{Column} \\ & & j & & \end{array} \\
\begin{array}{c} \\ \\ \text{Row} \\ i \\ \\ \end{array} & \begin{bmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{i,1} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \vdots & & \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,j} & \cdots & a_{m,n} \end{bmatrix} \\
 & \begin{array}{ccccc} \uparrow & & \uparrow & & \uparrow \\ \mathbf{a}_1 & & \mathbf{a}_j & & \mathbf{a}_n \end{array}
\end{array}$$

Example 2.2.1 Nice 2×2-matrices

Identity matrix:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Zero matrix:

$$\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Symmetric and **anti-symmetric** matrices:

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}, \begin{bmatrix} 0 & d \\ -d & 0 \end{bmatrix}$$

for numbers a, b, c, d.

Upper- and lower triangular matrices:

$$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}, \begin{bmatrix} d & 0 \\ e & f \end{bmatrix}$$

for numbers a, b, c, d, e, f.

Question 2.2.1

Can you write down 3×3 (anti-)symmetric and upper/lower triangular matrices?

Let us look at some uglier matrices as well.

Example 2.2.2

$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \end{bmatrix}$	2 rows and 3 columns	(2×3)-matrix
$B = \begin{bmatrix} 3 & 8 \\ 9 & 8 \\ 0 & 1 \end{bmatrix}$	3 rows and 2 columns	(3×2)-matrix
$C = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 4 & 1 \end{bmatrix}$	3 rows and 3 columns	(3×3)-matrix
$D = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}$	3 rows and 1 columns	(3×1)-matrix
$E = \begin{bmatrix} 1 & 3 & 2 \end{bmatrix}$	1 rows and 3 columns	(1×3)-matrix

Note that the matrix D and the matrix E can also be seen as vectors. Hence we give the following alternative definition of vectors.

Definition 2.2.2 Alternative vector definition

A **vector** is a matrix having either only one row or one column. Matrices having only one column ($n \times 1$-matrices) are called **column vectors** and matrices having only one row $(1 \times n)$ are called **row vectors**.

Theorem 2.2.1

Let A be an $m \times n$-matrix, $\alpha, \beta \in \mathbb{R}$ a number and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ vectors, then

$$A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha A\mathbf{x} + \beta A\mathbf{y}$$

Equivalently, the function $T_A \colon \mathbb{R}^n \to \mathbb{R}^m$ given by

$$T_A(\mathbf{x}) = A\mathbf{x}$$

is a linear transformation.

That is, **Any mapping defined by a matrix is a linear transformation.**

Actually, we have that the linear transformations described this way gives us every linear transformation.

Theorem 2.2.2

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. There exists a unique $m \times n$-matrix A such that

$$T(\mathbf{x}) = A\mathbf{x}, \qquad (\mathbf{x} \in \mathbb{R}^n).$$

This matrix is given by

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$$

where $\mathbf{e}_i$ is **the ith standard basis vector** of $\mathbf{R}^n$, which consists of a 1 in the ith row and zero elsewhere.

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Example 2.2.3 Matrices for some Linear Transformations

Consider the linear transformations from Example 2.1.4. We will rewrite these using a matrix.

- $\mathrm{Id} : \mathbb{R}^2 \to \mathbb{R}^2$ has the matrix $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,
- $S : \mathbb{R}^2 \to \mathbb{R}^2, S(x) = Ax$ has the matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$,
- $T : \mathbb{R}^3 \to \mathbb{R}^3, T(x) = bx$ has the matrix $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

2.3 Matrix operations

Addition and subtraction

As with numbers, we can add and subtract matrices. This is done element wise and only if the matrices are of the same size.

Example 2.3.1

$$\begin{bmatrix} 1 & 2 & 3 \\ 7 & 3 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 5 & 1 \\ 9 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1+2 & 2+5 & 3+1 \\ 7+9 & 3+3 & 2+0 \end{bmatrix} = \begin{bmatrix} 3 & 7 & 4 \\ 16 & 6 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 3 \\ 7 & 1 \\ 5 & 9 \end{bmatrix} - \begin{bmatrix} 2 & 8 \\ 3 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -2 & -5 \\ 4 & 0 \\ 5 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 7 & 3 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 3 \\ 7 & 1 \\ 5 & 9 \end{bmatrix} = \text{undefined}$$

Addition on numbers is commutative, meaning that the order of the numbers are indifferent.

$$34 + 79 = 113$$
$$79 + 34 = 113$$

There is also a special number (additive identity) under addition which has no effect on the other, namely *zero*.

$$340 + 0 = 340$$

$$0 + 59 = 59$$

Both of these properties have their analogue for addition in matrices, since addition on those are defined element wise.

Theorem 2.3.1

Let A, B and C be matrices of equal size $m \times n$, then

- $A + B = B + A$ (Additive commutativity)
- $(A + B) + C = A + (B + C)$ (Additive associativity)
- $A + \mathbf{0} = A$ (Additive identity)

where $\mathbf{0}$ is the zero matrix of size $m \times n$, that is the matrix with all elements equal zero.

Scalar multiplication

If we have a number c and a matrix A, then we can *scale* the matrix using the operation **scalar multiplication**. This is done by taking the produkt of c element wise on the matrix A, and we denote the resulting matrix with cA.

$$cA = c\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} = \begin{bmatrix} ca_{1,1} & ca_{1,2} & \cdots & ca_{1,n} \\ ca_{2,1} & ca_{2,2} & \cdots & ca_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m,1} & ca_{m,2} & \cdots & ca_{m,n} \end{bmatrix}$$

Example 2.3.2

Let us look at the scalar multiplication of the matrices from example 2.2.2

$$3A = 3\begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ 0 & 9 & 3 \end{bmatrix}$$

$$4B = 4\begin{bmatrix} 3 & 8 \\ 9 & 8 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 12 & 32 \\ 36 & 32 \\ 0 & 4 \end{bmatrix}$$

$$2C = 2\begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 0 \\ 4 & 6 & 2 \\ 0 & 8 & 2 \end{bmatrix}$$

$$5D = 5\begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \\ 20 \end{bmatrix}$$

$$8E = 8\begin{bmatrix} 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 24 & 16 \end{bmatrix}$$

This operation also inherit some good looking properties from multiplication of numbers. Recall that when you multiply numbers with an addition of numbers in parentheses you can either sum up the addition and then multiply or multiply each summand and then add.

$$3(4+2) = \begin{cases} 3 \cdot 6 = 18 \\ 3 \cdot 4 + 3 \cdot 2 = 12 + 6 = 18 \end{cases}$$

$$(5+1)4 = \begin{cases} 6 \cdot 4 = 24 \\ 5 \cdot 4 + 1 \cdot 4 = 20 + 4 = 24 \end{cases}$$

Theorem 2.3.2

Let A and B be matrices of equal size $m \times n$ and c, d be two numbers, then

- $c(dA) = (cd)A$

- $c(A+B) = cA + cB$
- $(c+d)A = cA + dA$
- $1A = A$

Multiplication of matrices

Another operation of numbers which we would like to emulate for matrices are multiplication. There is a naivé way of defining such an operation by doing it element wise. This turns out to be of little to no benefit, so let us disregard that idea at once.

Let us recall that linear transformations are uniquely determined by a matrix. Therefore, when multiplying a vector $\mathbf{x} \in \mathbb{R}^n$ with a $m \times n$-matrix, we can think of it as a function from $\mathbb{R}^n$ to $\mathbb{R}^m$. We want to think of multiplication of two matrices as the composition of two function, and therefore the following equality has to hold:

$$(AB)\mathbf{x} = A(B\mathbf{x})$$

The linear transformation T_B (multiplication by B) acts first and then the transformation T_A (multiplication by A). If A is a $m \times n$-matrix and B a $q \times p$-matrix, then we need $n = p$ for this to make sense, since T_B sends vectors in $\mathbb{R}^p$ to vectors in $\mathbb{R}^q$ and T_A sends vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$.

$$\begin{tikzcd} \mathbb{R}^q \arrow[r, "T_B"] \arrow[rr, bend right, "T_{AB}"'] & \mathbb{R}^p = \mathbb{R}^n \arrow[r, "T_A"] & \mathbb{R}^m \end{tikzcd}$$

The composition of T_B with T_A will be T_{AB}. Before defining what AB has to be for this to make sense, let us look at an example.

Example 2.3.3

Let A and B be the following two 2×2-matrices:

$$A = \begin{bmatrix} 3 & -5 \\ 2 & 7 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$$

We want to find the matrix AB such that $(AB)\mathbf{v} = A(B\mathbf{v})$ for all vectors $\mathbf{v}$ i $\mathbb{R}^2$.

We first look at multiplication with the standard basis vectors

$$\begin{bmatrix}1\\0\end{bmatrix} \quad \text{og} \quad \begin{bmatrix}0\\1\end{bmatrix}.$$

We know from the definition of matrix multiplication on vectors that by multiplying a 2×2-matrix with one of these vectors, we obtain either the first or second column of the matrix.

$$B\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}4\\2\end{bmatrix}$$
$$A\left(B\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}3 & -5\\2 & 7\end{bmatrix}\begin{bmatrix}4\\2\end{bmatrix} = \begin{bmatrix}2\\22\end{bmatrix}$$

We want AB to be given in a way that satisfies

$$(AB)\mathbf{v} = A(B\mathbf{v})$$

for all vectors $\mathbf{v}$. In particular:

$$(AB)\begin{bmatrix}1\\0\end{bmatrix} = A\left(B\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}2\\22\end{bmatrix}$$

Since multiplication with

$$\begin{bmatrix}1\\0\end{bmatrix}$$

gives us the first column of the matrix, we know that the first column of AB has to be

$$\begin{bmatrix}2\\22\end{bmatrix}$$

Likewise the second of column of AB is:

$$B\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}3\\1\end{bmatrix}$$
$$A\left(B\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}3 & -5\\2 & 7\end{bmatrix}\begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}4\\13\end{bmatrix}$$
$$(AB)\begin{bmatrix}0\\1\end{bmatrix} = A\left(B\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}4\\13\end{bmatrix}$$

Then we see that the second column of AB needs to be:

$$\begin{bmatrix} 4 \\ 13 \end{bmatrix}$$

So the product of A og B is:

$$AB = \begin{bmatrix} 2 & 4 \\ 22 & 13 \end{bmatrix}$$

In general if A is of size $m \times n$, B is of size $n \times p$ and $\mathbf{x} \in \mathbb{R}^p$, then

$$B\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_p\mathbf{b}_p$$

for $\mathbf{b}_1, \dots, \mathbf{b}_p$ the columns of B and $x_1, \dots, x_n$ the elements of $\mathbf{x}$. Now using that $A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha A\mathbf{x} + \beta A\mathbf{y}$, we have

$$A(B\mathbf{x}) = x_1 A\mathbf{b}_1 + x_2 A\mathbf{x}_2 + \cdots + x_p A\mathbf{b}_p$$

We recognize this as being equal to

$$\begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix} \mathbf{x}$$

and we conclude that $AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}$, or using the definition of multiplication of matrix with vector, we have

$$AB = \begin{bmatrix} \mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 & \cdots & \mathbf{a}_1\mathbf{b}_p \\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 & \cdots & \mathbf{a}_2\mathbf{b}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m\mathbf{b}_1 & \mathbf{a}_m\mathbf{b}_2 & \cdots & \mathbf{a}_m\mathbf{b}_p \end{bmatrix}$$

where $\mathbf{a}_1, \dots \mathbf{a}_n$ is the row vectors of A.

Definition 2.3.1

Let A be an $m \times n$-matrix and B an $n \times p$-matrix. Their product AB is an $m \times p$ matrix where the element $(AB)_{i,j}$ is

$$(AB)_{i,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots + a_{i,n}b_{n,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j}$$

which is the same as saying that $(AB)_{i,j}$ is the dot product of the ith row vector of A and the jth column vector of B.

$$m \overbrace{\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ a_{i,1} & a_{i,2} & \cdots & a_{i,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}}^{n} \times n \overbrace{\begin{bmatrix} b_{1,1} & \cdots & b_{1,j} & \cdots & b_{1,p} \\ b_{2,1} & \cdots & b_{2,j} & \cdots & b_{2,p} \\ \vdots & & \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,j} & \cdots & b_{n,p} \end{bmatrix}}^{p} = m \overbrace{\begin{bmatrix} ab_{1,1} & \cdots & ab_{1,j} & \cdots & ab_{1,p} \\ \vdots & & \vdots & & \vdots \\ ab_{i,1} & \cdots & ab_{i,j} & \cdots & ab_{i,p} \\ \vdots & & \vdots & & \vdots \\ ab_{m,1} & \cdots & ab_{m,j} & \cdots & ab_{m,p} \end{bmatrix}}^{p}$$

$$\text{row } i \qquad \times \qquad \text{column } j \qquad = \qquad (i,j)\text{-element}$$

Example 2.3.4

A possibly easier way of visualizing the process is the following.

Say we want to calculate the following product

$$\overbrace{\begin{bmatrix} 1 & 2 & 3 \\ 7 & 3 & 2 \end{bmatrix}}^{2\times 3} \overbrace{\begin{bmatrix} 0 & 3 \\ 7 & 1 \\ 5 & 9 \end{bmatrix}}^{3\times 2} = \overbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}^{2\times 2}$$

then we can organize the matrices as following

$$\begin{array}{ccc|cc} & & & 0 & 3 \\ & & & 7 & 1 \\ & & & 5 & 9 \\ \hline 1 & 2 & 3 & a & b \\ 7 & 3 & 2 & c & d \end{array}$$

The second matrix is placed in the upper right corner and the first matrix in the lower left corner. In order to find the elements in the product, we take the dot product of the row and column pointing in to the element.

$$\begin{array}{ccc|cc} & & & 0 & 3 \\ & & & 7 & 1 \\ & & & 5 & 9 \\ \hline 1 & 2 & 3 & a & b \\ 7 & 3 & 2 & c & d \end{array}$$

$$a = 1 \cdot 0 + 2 \cdot 7 + 3 \cdot 5 = 29$$

$$\begin{array}{ccc|cc} & & & 0 & 3 \\ & & & 7 & 1 \\ & & & 5 & 9 \\ \hline 1 & 2 & 3 & a & b \\ 7 & 3 & 2 & c & d \end{array}$$

$$b = 1 \cdot 3 + 2 \cdot 1 + 3 \cdot 9 = 32$$

This gives us

$$\begin{bmatrix} 1 & 2 & 3 \\ 7 & 3 & 2 \end{bmatrix} \begin{bmatrix} 0 & 3 \\ 7 & 1 \\ 5 & 9 \end{bmatrix} = \begin{bmatrix} 29 & 32 \\ 31 & 42 \end{bmatrix}$$

Remark 2.3.1

Be aware that matrix multiplication is only defined if the matrices has the right corresponding sizes.

$$[m\times \cancel{n}] \cdot [\cancel{n} \times p] = m \times p$$

Example 2.3.5

a) Calculate, if possible,

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

$$\begin{array}{ccc|ccc} & & & 1 & 1 & 0 \\ & & & 2 & 3 & 1 \\ & & & 0 & 4 & 1 \\ \hline 1 & 2 & 3 & 5 & 19 & 5 \\ 7 & 3 & 2 & 13 & 24 & 5 \end{array}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 7 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \underline{\begin{bmatrix} 5 & 19 & 5 \\ 13 & 24 & 5 \end{bmatrix}}$$

b) Calculate, if possible,

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

This is undefined since the number of columns in the first matrix do not equal the number of rows in the second matrix.

We have seen that additive properties of numbers have corresponding properties for addition of matrices. The same holds true for multiplication, except for one crucial detail, namely commutativity, i.e. $a \times b = b \times a$. We have already seen an extreme way it fails to commute in the preceding example, namely because of size constraints, but as we will now see it is generally not true even for square matrices $(n \times n)$.

Example 2.3.6 Non-commutivity

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

However, we have a lot of other nice properties of matrix-multiplication.

Theorem 2.3.3

Let A, B and C be matrices such that the operations described below is defined, and d a number. Then

- Multiplicative associativity

$$A(BC) = (AB)C$$

- Multiplicative distributivity

$$A(B+C) = AB + AC \text{ and } (A+B)C = AC + BC$$

- $(dA)B = d(AB) = A(dB)$

Let now I_n denote a $n \times n$-matrix with 1's on the diagonal and 0's everywhere else, then for A a $m \times n$-matrix we have

- $AI_n = A = I_m A$. Thus I_n is called the $n \times n$ identity matrix.

Remark 2.3.2

Let us stress the fact that for two matrices A and B, we do in general ***not*** have

$$AB = BA$$

Example 2.3.7

Let A, B and C be the following matrices:

$$A = \begin{bmatrix} 5 & 0 & -2 \\ 3 & 1 & 4 \end{bmatrix}$$

$$B = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 2 & 4 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

A is of size 2×3 and B is of size 3×2, so AB must be of size 2×2. Let us calculate this using the definition:

$$\begin{aligned} AB &= \begin{bmatrix} 5 \cdot 2 + 0 \cdot 1 + (-2) \cdot 2 & 5 \cdot 1 + 0 \cdot 0 + (-2) \cdot 4 \\ 3 \cdot 2 + 1 \cdot 1 + 4 \cdot 2 & 3 \cdot 1 + 1 \cdot 0 + 4 \cdot 4 \end{bmatrix} \\ &= \begin{bmatrix} 6 & -3 \\ 15 & 19 \end{bmatrix} \end{aligned}$$

Multiplying the same matrices, but in opposite order gives a 3×3-matrix:

$$
\begin{aligned}
BA &= \begin{bmatrix} 2{\cdot}5+1{\cdot}3 & 2{\cdot}0+1{\cdot}1 & 2{\cdot}(-2)+1{\cdot}4 \\ 1{\cdot}5+0{\cdot}3 & 1{\cdot}0+0{\cdot}1 & 1{\cdot}(-2)+0{\cdot}4 \\ 2{\cdot}5+4{\cdot}3 & 2{\cdot}0+4{\cdot}1 & 2{\cdot}(-2)+4{\cdot}4 \end{bmatrix} \\
&= \begin{bmatrix} 13 & 1 & 0 \\ 5 & 0 & -2 \\ 22 & 4 & 12 \end{bmatrix}
\end{aligned}
$$

The product of B and C is the 3×2-matrix

$$BC = \begin{bmatrix} 5 & 8 \\ 1 & 2 \\ 14 & 20 \end{bmatrix}$$

However the product CB is not defined, since C is of size 2×2 and B is of size 3×2.

We could have calculated CA, but AC is also not defined.

Remark 2.3.3

Another important remark is that cancellation of matrices are in general not possible. That is, if $AB = AC$, then we do **not** necessarily have $B = C$. For example

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

However,

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Matrix transpose

Before moving on to the more interesting concept of inverse matrices, let us as an interlude mention a matrix operation which do not have an equal in number manipulation.

Definition 2.3.2

Let

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

be a $m \times n$-matrix. The **transpose** of A is the $n \times m$-matrix

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

where the rows are the columns of A and the columns are the rows in A.

A square matrix A is **symmetric** if and only if $A = A^T$.

Example 2.3.8

If we let A be the matrix

$$A = \begin{bmatrix} 5 & 0 & -2 \\ 3 & 1 & 4 \end{bmatrix},$$

then the transpose of A is:

$$A^T = \begin{bmatrix} 5 & 3 \\ 0 & 1 \\ -2 & 4 \end{bmatrix}.$$

Theorem 2.3.4

For every matrix A we have:

$$(A^T)^T = A.$$

If A and B are matrices such that the product AB is defined, then:

$$(AB)^T = B^T \cdot A^T.$$

2.4 Invertability

For every number a not equal to zero, there exist an inverse number $a^{-1} = \frac{1}{a}$ such that $a \cdot a^{-1} = 1 = a^{-1}a$. It is natural to wonder whether this property extends to matrices. Let us limit the discussion to only square matrices, i.e. matrices of size $n \times n$ for some n. The question is then: Given a square $n \times n$-matrix A, can we find a matrix B such that the following equalities hold?

$$AB = I_n = BA.$$

Let us look at an example of such matrices.

Example 2.4.1

Let A and B be the following 2×2-matrices:

$$A = \begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix} \quad \text{og} \quad B = \begin{bmatrix} -1/2 & 1 \\ -3/2 & 2 \end{bmatrix}$$

Then we have

$$A \cdot B = \begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} -1/2 & 1 \\ -3/2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

and

$$B \cdot A = \begin{bmatrix} -1/2 & 1 \\ -3/2 & 2 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

These matrices therefore satisfies

$$A \cdot B = I_2 = B \cdot A.$$

Definition 2.4.1

Let A be a matrix of size $n \times n$. An **inverse** of A is a $n \times n$-matrix B such that

$$AB = I_n = BA.$$

A matrix is **invertible** if it has an inverse.

A square matrix do not necessarily have an inverse. The following three matrices are examples of such.

Example 2.4.2

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

are not invertible matrices.

Theorem 2.4.1

If a matrix has an inverse, then the inverse matrix is unique. We denote this matrix by A^{-1}.

One reason to be interested in the invertability of matrices comes from the solvability of linear equations. Remember that if we have a system of n linear equations in n variables, then we can construct the corresponding matrix equation $A\mathbf{x} = \mathbf{b}$ where A is the coefficient matrix.

Theorem 2.4.2

Let A be a $n \times n$-matrix, and $\mathbf{b} \in \mathbb{R}^n$ a vector. If A is invertible then $A\mathbf{x} = \mathbf{b}$ has a solution. The solution is unique and given by $A^{-1}\mathbf{b}$

Moreover, if $A\mathbf{x} = \mathbf{c}$ has a solution for every $\mathbf{c} \in \mathbb{R}^n$, then A is invertible.

Simultaneous solutions of multiple systems

Assume that we want to solve multiple matrix equations

$$\begin{aligned} A\mathbf{x}_1 &= \mathbf{b}_1 \\ A\mathbf{x}_2 &= \mathbf{b}_2 \\ &\vdots \\ A\mathbf{x}_t &= \mathbf{b}_t \end{aligned}$$

with the same coefficient matriz A, but different vectors on the right $\mathbf{b}_1$, $\mathbf{b}_2$, ..., $\mathbf{b}_t$. Then we could reduce the augmented matrix of each system individually:

$$\begin{gathered} \left[A \mid \mathbf{b}_1\right] \\ \left[A \mid \mathbf{b}_2\right] \\ \vdots \\ \left[A \mid \mathbf{b}_t\right] \end{gathered}$$

Or we could observe that since the left hand side is equal for each of them, we are doing the same thing multiple times except for what is on the right side. We can save ourselves a lot of trouble if we combine all the augmented matrices into one:

$$\left[\begin{array}{c|cccc} A & \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_t \end{array}\right],$$

and reduce that instead.

Example 2.4.3

Let A be the following matrix:

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

We would like to solve the following systems:

$$A\mathbf{x}_1 = \begin{bmatrix} -2 \\ 7 \end{bmatrix} \qquad A\mathbf{x}_2 = \begin{bmatrix} 10 \\ 1 \end{bmatrix} \qquad A\mathbf{x}_3 = \begin{bmatrix} -4 \\ 0 \end{bmatrix}$$

We combine them all into a big augmented matrix and reduce

$$\begin{aligned}
\left[\begin{array}{cc|ccc} 2 & -2 & -2 & 10 & -4 \\ 1 & 3 & 7 & 1 & 0 \end{array}\right] &\sim \left[\begin{array}{cc|ccc} 1 & 3 & 7 & 1 & 0 \\ 2 & -2 & -2 & 10 & -4 \end{array}\right] \\
&\sim \left[\begin{array}{cc|ccc} 1 & 3 & 7 & 1 & 0 \\ 0 & -8 & -16 & 8 & -4 \end{array}\right] \\
&\sim \left[\begin{array}{cc|ccc} 1 & 3 & 7 & 1 & 0 \\ 0 & 1 & 2 & -1 & 1/2 \end{array}\right] \\
&\sim \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 4 & -3/2 \\ 0 & 1 & 2 & -1 & 1/2 \end{array}\right]
\end{aligned}$$

The last matrix is in reduced row echelon form, and we read the solutions of the systems from the right hand side of this matrix.

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 4 \\ -1 \end{bmatrix} \qquad \mathbf{x}_3 = \begin{bmatrix} -3/2 \\ 1/2 \end{bmatrix}$$

Finding the inverse matrix

Now, let us find a formula for the inverse matrices. Assume that we have a $n \times n$-matrix A. We would ascertain whether it is invertible and if so find the inverse matrix A^{-1}.

We consider the equation

$$AX = I_n,$$

where X is an unknown $n \times n$-matrix. Let

$$\mathbf{x}_1, \ \mathbf{x}_2, \ \ldots, \ \mathbf{x}_n$$

be the columns of X, that is

$$X = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix},$$

then we can write the product AX as

$$AX = \begin{bmatrix} A\mathbf{x}_1 & A\mathbf{x}_2 & \cdots & A\mathbf{x}_n \end{bmatrix}$$

Let us name the columns of I_n as well:

$$I_n = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \ldots & \mathbf{e}_n \end{bmatrix}$$

That means that $\mathbf{e}_i$ is the vector in $\mathbb{R}^n$ with 1 in the ith row, and 0's everywhere else.

The equation $AX = I_n$ is now equal to the equations:

$$\begin{aligned} A\mathbf{x}_1 &= \mathbf{e}_1 \\ A\mathbf{x}_2 &= \mathbf{e}_2 \\ &\vdots \\ A\mathbf{x}_n &= \mathbf{e}_n \end{aligned}$$

By the technique described above we can therefore reduce the augmented matrix

$$\left[\begin{array}{c|cccc} A & \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{array} \right] = \left[\begin{array}{c|c} A & I_n \end{array} \right]$$

to solve for X.

Example 2.4.4

We want to invert the following matrix if possible:

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

Following the ideas above, we would like to find a matrix X such that $AX = I_3$

by solving the three systems:

$$A\mathbf{x}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \quad A\mathbf{x}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix} \quad A\mathbf{x}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

We make the combined augmented matrix for the systems and reduce:

$$\left[\begin{array}{ccc|ccc}1 & -1 & -3 & 1 & 0 & 0\\0 & 1 & 3 & 0 & 1 & 0\\2 & -2 & -1 & 0 & 0 & 1\end{array}\right] \sim \left[\begin{array}{ccc|ccc}1 & -1 & -3 & 1 & 0 & 0\\0 & 1 & 3 & 0 & 1 & 0\\0 & 0 & 5 & -2 & 0 & 1\end{array}\right]$$
$$\sim \left[\begin{array}{ccc|ccc}1 & -1 & -3 & 1 & 0 & 0\\0 & 1 & 3 & 0 & 1 & 0\\0 & 0 & 1 & -\frac{2}{5} & 0 & \frac{1}{5}\end{array}\right]$$
$$\sim \left[\begin{array}{ccc|ccc}1 & 0 & 0 & 1 & 1 & 0\\0 & 1 & 3 & 0 & 1 & 0\\0 & 0 & 1 & -\frac{2}{5} & 0 & \frac{1}{5}\end{array}\right]$$
$$\sim \left[\begin{array}{ccc|ccc}1 & 0 & 0 & 1 & 1 & 0\\0 & 1 & 0 & \frac{6}{5} & 1 & -\frac{3}{5}\\0 & 0 & 1 & -\frac{2}{5} & 0 & \frac{1}{5}\end{array}\right]$$

We obtain from this the following solutions:

$$\mathbf{x}_1 = \begin{bmatrix}1\\6/5\\-2/5\end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix}1\\1\\0\end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix}0\\-3/5\\1/5\end{bmatrix}$$

Now, X is given by

$$X = \begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3\end{bmatrix} = \begin{bmatrix}1 & 1 & 0\\6/5 & 1 & -3/5\\-2/5 & 0 & 1/5\end{bmatrix}$$

We have found X by solving $AX = I_3$, and if you try taking the product XA, you will find that this also equals I_3. We conclude that X is the inverse of A and A is invertible, that is $X = A^{-1}$

In the example we solved the equation $AX = I_3$, and it turned out that X also solved $XA = I_3$, so we concluded that $A^{-1} = X$. This actually holds in general.

Theorem 2.4.3

Let A be a $n \times n$-matrix.

(a) A is invertible if and only if the reduced row echelon form of A equals I_n.

(b) If A is invertible, we can find the inverse by reducing the matrix

$$\left[A \mid I_n\right]$$

to reduced row echelon form and read out the right hand side. In other word: The reduced row echelon form is on the form:

$$\left[I_n \mid A^{-1}\right]$$

Formula for inverting a 2×2-matrix

For a 2×2-matrix, there is an explicit formula for the inverse.

Theorem 2.4.4

Let

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

be invertible. Then

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

Proof.

$$AA^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

The product $A^{-1}A = I_2$ is calculated in the same way. $\square$