Sample problem — a system of linear equations

Solve the following system of linear equations:

$$\begin{aligned} x_1 + 4x_2 + 3x_3 + 2x_4 + x_5 &= 1 \\ 2x_1 + 8x_2 + 4x_3 \phantom{{}+ 6x_4 + 9x_5} &= 0 \\ 3x_3 + 6x_4 + 9x_5 &= 5 \\ 2x_1 + 8x_2 + 7x_3 + 6x_4 + 3x_5 &= 3 \end{aligned}$$

Questions:

- How to efficiently describe the system?
- What do we mean by a solution of the system?
- How to get some or all solutions of the system?

Real vectors

Definition: A real *vector* $\boldsymbol{b}$ with m components is an ordered m-tuple of real numbers $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$. We write $\boldsymbol{b} \in \mathbb{R}^m$.

We consider *column* vectors.
For the row-wise notation
we use the transposition, i.e. $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = (b_1, b_2, \ldots, b_m)^\mathsf{T}$.

The vector $\mathbf{0}_m = (0, \ldots, 0)^\mathsf{T} \in \mathbb{R}^m$ is the *zero vector*.
When the context is clear, it can be written shortly as $\mathbf{0}$.

An ordered n-tuple of variables $\boldsymbol{x} = (x_1, \ldots, x_n)^\mathsf{T}$
is the *vector of unknowns*.

Real matrices

Definition: A real $m \times n$ *matrix* $\boldsymbol{A}$ is a collection of mn real numbers arranged in an array with m rows and n columns.

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

We write $\boldsymbol{A} \in \mathbb{R}^{m \times n}$.

The elements of a matrix are denoted as $(\boldsymbol{A})_{i,j} = a_{i,j}$.
If the indices are obvious, the comma can be omitted and only a_{ij} or $(\boldsymbol{A})_{ij}$ can be written.
Otherwise, we leave the comma, e.g., to distinguish $a_{12,3}$ from $a_{1,23}$; or $a_{i,jk}$ from $a_{ij,k}$; or for the element $a_{i,j(i)}$, etc.

A *square matrix* of *order* n has n rows and n columns.

Systems of linear equations

Definition: Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{b} \in \mathbb{R}^m$ and $\boldsymbol{x} = (x_1, \dots, x_n)^\mathsf{T}$ is a vector of unknowns.

The *system of m linear equations in n unknowns* is $\boldsymbol{Ax} = \boldsymbol{b}$, in expanded form written as:

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

The matrix $\boldsymbol{A}$ is the *matrix of the system*,
the vector $\boldsymbol{b}$ is the *right-hand side vector*,
the matrix $(\boldsymbol{A}|\boldsymbol{b}) \in \mathbb{R}^{m \times (n+1)}$ is the *augmented matrix*.

A vector $\boldsymbol{x} \in \mathbb{R}^n$ is a *solution* of the system $\boldsymbol{Ax} = \boldsymbol{b}$ if it satisfies all its m equations, i.e:

$$\forall i \in \{1, \dots, m\} : a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i.$$

Systems $\boldsymbol{Ax} = \mathbf{0}$ are called *homogeneous* and always allow $\boldsymbol{x} = \mathbf{0}$.

Example

The augmented matrix of the system of linear equations

$$\begin{aligned} x_1 + 4x_2 + 3x_3 + 2x_4 + x_5 &= 1 \\ 2x_1 + 8x_2 + 4x_3 &= 0 \\ 3x_3 + 6x_4 + 9x_5 &= 5 \\ 2x_1 + 8x_2 + 7x_3 + 6x_4 + 3x_5 &= 3 \end{aligned}$$

is formed by its coefficients:

$$(\boldsymbol{A}|\boldsymbol{b}) = \left(\begin{array}{ccccc|c} 1 & 4 & 3 & 2 & 1 & 1 \\ 2 & 8 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 6 & 9 & 5 \\ 2 & 8 & 7 & 6 & 3 & 3 \end{array}\right)$$

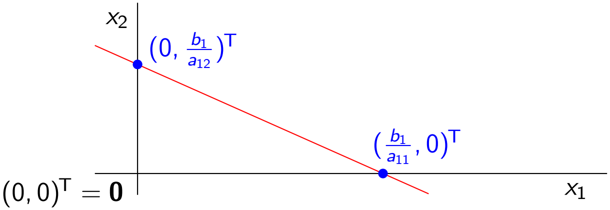
The vector $\boldsymbol{x} = (x_1, \ldots, x_5)^\mathsf{T} = \left(4, -1, 0, \frac{1}{3}, \frac{1}{3}\right)^\mathsf{T}$ is a possible solution of this system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, since it satisfies all its equations:

$$\begin{aligned} 4 + 4 \cdot (-1) + 3 \cdot 0 + 2 \cdot \tfrac{1}{3} + \tfrac{1}{3} &= 1 \\ 2 \cdot 4 + 8 \cdot (-1) + 4 \cdot 0 \phantom{+ 6 \cdot \tfrac{1}{3} + 9 \cdot \tfrac{1}{3}} &= 0 \\ 3 \cdot 0 + 6 \cdot \tfrac{1}{3} + 9 \cdot \tfrac{1}{3} &= 5 \\ 2 \cdot 4 + 8 \cdot (-1) + 7 \cdot 0 + 6 \cdot \tfrac{1}{3} + 3 \cdot \tfrac{1}{3} &= 3 \end{aligned}$$

Geometric meaning — one equation in two unknowns

$a_1x_1 + a_{12}x_2 = b_1$ or equivalently $\begin{pmatrix} a_{11} & a_{12} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \end{pmatrix}$

- If $a_{11} \neq 0$ or $a_{12} \neq 0$ then the set of solutions forms a line in the Euclidean plane.



- It could be parallel to one of the axes, e.g. to x_1, if $a_{11} = 0$.

Degenerate cases:

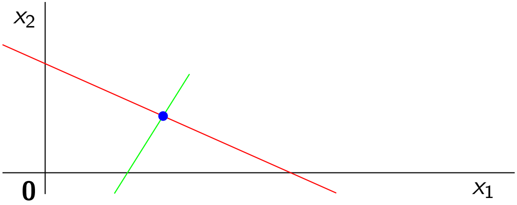
- If $a_{11} = a_{12} = 0$ and $b_1 \neq 0$, then the system has no solution.
- If $a_{11} = a_{12} = 0$ and $b_1 = 0$, then all points of the Euclidean plane are solutions.

Two equations in two unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \qquad \text{or} \qquad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

If both equations are nondegenerate, then the set of solutions is the intersection of two lines, which could be

- a point:



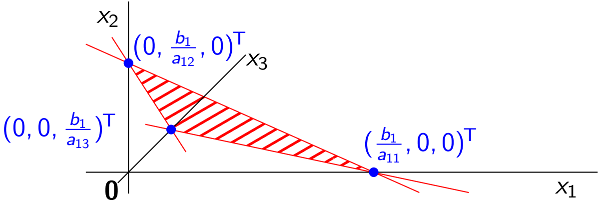
- an empty set, if the two lines are distinct parallel,
- a line, if the two lines are identical.

Among the degenerate cases, the system $\mathbf{0x} = \mathbf{0}$
also yields all points of the Euclidean plane as the set of solutions.

One equation in three unknowns

$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$

- In the nondegenerate case $a_{11} \neq 0 \vee a_{12} \neq 0 \vee a_{13} \neq 0$ solutions form a plane in the 3-dimensional Euclidean space:



Degenerate cases:

- If $a_{11} = a_{12} = a_{13} = 0$ and $b_1 \neq 0$, then no solution exists.
- If $a_{11} = a_{12} = a_{13} = 0$ and $b_1 = 0$ then all points of the Euclidean space are solutions.

Elementary equivalent row transformations

Definition: We write $\boldsymbol{A} \sim \boldsymbol{A}'$ if $\boldsymbol{A}'$ is obtained from $\boldsymbol{A}$ by any of the following *elementary equivalent row transformations*:

1. Multiplication of the i-th row by a *nonzero* $t \in \mathbb{R} \setminus \{0\}$,

   formally: $a'_{kl} = \begin{cases} a_{kl} & \text{if } k \neq i \\ ta_{il} & \text{if } k = i \end{cases}$

2. Adding the j-th row to the i-th row,

   formally: $a'_{kl} = \begin{cases} a_{kl} & \text{if } k \neq i \\ a_{il} + a_{jl} & \text{if } k = i \end{cases}$

From the two above the following two can be derived:

3. Adding the j-th row multiplied by $t \in \mathbb{R}$ to the i-th row.
4. Exchange of two rows.

A series of elementary transformations is denoted as $\boldsymbol{A} \sim\sim \boldsymbol{A}'$.

Use of elementary transformations

Theorem: Let $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$ be two systems such that $(\boldsymbol{A}|\boldsymbol{b}) \sim\sim (\boldsymbol{A}'|\boldsymbol{b}')$. Then both systems have identical solution sets.

Example:

$$(\boldsymbol{A}|\boldsymbol{b}) = \left(\begin{array}{ccccc|c} 1 & 4 & 3 & 2 & 1 & 1 \\ 2 & 8 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 6 & 9 & 5 \\ 2 & 8 & 7 & 6 & 3 & 3 \end{array}\right) \begin{array}{c} \cdot 2 \\ \\ \\ \\ \end{array} \sim \left(\begin{array}{ccccc|c} 2 & 8 & 6 & 4 & 2 & 2 \\ 2 & 8 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 6 & 9 & 5 \\ 2 & 8 & 7 & 6 & 3 & 3 \end{array}\right) \begin{array}{c} -\mathrm{II} \\ \\ \\ \\ \end{array} \sim$$

$$\left(\begin{array}{ccccc|c} 0 & 0 & 2 & 4 & 2 & 2 \\ 2 & 8 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 6 & 9 & 5 \\ 2 & 8 & 7 & 6 & 3 & 3 \end{array}\right) \begin{array}{c} \\ \\ \mathrm{IV} \\ \mathrm{III} \end{array} \sim \left(\begin{array}{ccccc|c} 0 & 0 & 2 & 4 & 2 & 2 \\ 2 & 8 & 4 & 0 & 0 & 0 \\ 2 & 8 & 7 & 6 & 3 & 3 \\ 0 & 0 & 3 & 6 & 9 & 5 \end{array}\right) = (\boldsymbol{A}'|\boldsymbol{b}')$$

The vector $\boldsymbol{x} = \left(4, -1, 0, \frac{1}{3}, \frac{1}{3}\right)^{\mathsf{T}}$ solves also the system $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$, since $2 \cdot 0 + 4 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} = 2$, and the rest is only ordered differently.

However, the theorem applies not only to this particular solution and the chosen transformations, but also for *any possible* solution and *any sequence* of transformations.

Use of elementary transformations

Theorem: Let $\boldsymbol{Ax} = \boldsymbol{b}$ and $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$ be two systems such that $(\boldsymbol{A}|\boldsymbol{b}) \sim\sim (\boldsymbol{A}'|\boldsymbol{b}')$. Then both systems have identical solution sets.

Proof: It suffices to show that the solution set is preserved if a single transform of the first or of the second type is performed.
We aim to show that $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{Ax} = \boldsymbol{b}\} = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'\}$.
The set equality follows from two inclusions $\subseteq$ and $\supseteq$, seen as implications a: $\boldsymbol{Ax} = \boldsymbol{b} \Rightarrow \boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$, and b: $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}' \Rightarrow \boldsymbol{Ax} = \boldsymbol{b}$.

1a. $\boldsymbol{Ax} = \boldsymbol{b} \Rightarrow \boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$ for the i-th row scaling by $t \neq 0$:
As only the i-th row/equation is changed, any solution of $\boldsymbol{Ax} = \boldsymbol{b}$ satisfies the unchanged equations of $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$.
It remains to verify the i-th equation of $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$.
From the left hand side to the right: $a'_{i1}x_1 + \cdots + a'_{in}x_n =$
$ta_{i1}x_1 + \cdots + ta_{in}x_n = t(a_{i1}x_1 + \cdots + a_{in}x_n) = tb_i = b'_i$
Used: $a'_{il} = ta_{il}$ (definition), $tc + td = t(c + d)$ (extraction), $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$ (assumption), $tb_i = b'_i$ (definition).
Green indicates the relationship between $(\boldsymbol{A}|\boldsymbol{b})$ and $(\boldsymbol{A}'|\boldsymbol{b}')$ i.e. the elementary transformation; red the assumption $\boldsymbol{Ax} = \boldsymbol{b}$.

Summary of cases a single elementary transformation of the first or of the second second type and the i-th equation:

1a. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Rightarrow \boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}' : a'_{i1}x_1 + \cdots + a'_{in}x_n =$
$ta_{i1}x_1 + \cdots + ta_{in}x_n = t(a_{i1}x_1 + \cdots + a_{in}x_n) = tb_i = b'_i$

1b. $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}' \Rightarrow \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} : a_{i1}x_1 + \cdots + a_{in}x_n =$
$\frac{1}{t}(ta_{i1}x_1 + \cdots + ta_{in}x_n) = \frac{1}{t}(a'_{i1}x_1 + \cdots + a'_{in}x_n) = \frac{1}{t}b'_i = \frac{1}{t}tb_i = b_i$

2a. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Rightarrow \boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}' : a'_{i1}x_1 + \cdots + a'_{in}x_n =$
$(a_{i1} + a_{j1})x_1 + \cdots + (a_{in} + a_{jn})x_n =$
$(a_{i1}x_1 + \cdots + a_{in}x_n) + (a_{j1}x_1 + \cdots + a_{jn}x_n) = b_i + b_j = b'_i$

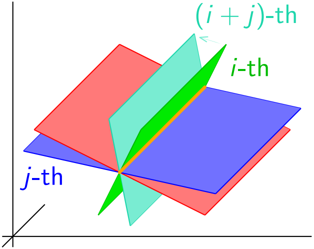
2b. $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}' \Rightarrow \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} : a_{i1}x_1 + \cdots + a_{in}x_n =$
$a_{i1}x_1 + \cdots + a_{in}x_n + b_j - b_j =$
$(a_{i1}x_1 + \cdots + a_{in}x_n) + (a_{j1}x_1 + \cdots + a_{jn}x_n) - b_j =$
$(a_{i1} + a_{j1})x_1 + \cdots + (a_{in} + a_{jn})x_n - b_j =$
$(a'_{i1}x_1 + \cdots + a'_{in}x_n) - b_j = b'_i - b_j = b_i + b_j - b_j = b_i$

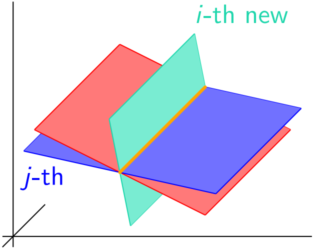
The color of $=$ means either the transform from $(\boldsymbol{A}|\boldsymbol{b})$ to $(\boldsymbol{A}'|\boldsymbol{b}')$, the assumption of the case, or an algebraic rearrangement of terms.

Geometric meaning of elementary transformations

1., 4. Multiplication of a row or swapping two rows does not change the position of the hyperplanes.

2., 3. Adding the j-th row (multiplied by $t \in \mathbb{R}$) to the i-th row moves the i-th hyperplane so that the intersection with the others remains unchanged.



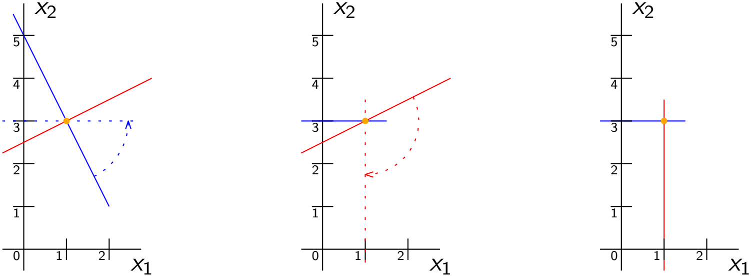


Geometric meaning of elementary transformations

1., 4. Multiplication of a row or swapping two rows does not change the position of the hyperplanes.

2., 3. Adding the j-th row (multiplied by $t \in \mathbb{R}$) to the i-th row moves the i-th hyperplane so that the intersection with the others remains unchanged.

Goal: move the hyperplanes so the solution could be seen easily.



$$2x_1 + x_2 = 5 \quad -x_1 + 2x_2 = 5 \quad (+2\times,\ :5)$$

$$x_2 = 3 \quad -x_1 + 2x_2 = 5 \quad (-2\times,\ \cdot(-1))$$

$$x_2 = 3 \quad x_1 = 1$$

Questions to understand the lecture topic

- How does the solution set of a system change geometrically when we alter $\boldsymbol{b}$, i.e. the right-hand side vector?
- How the systems without a solution in $\mathbb{R}^2$ and $\mathbb{R}^3$ can be described geometrically?
  (Discuss possible cases of positions of lines and planes.)
- May $t = 0$ in the third elementary transformation?
- Where was used the assumption $t \neq 0$ in the first elementary equivalent transformation?
- Which algebraic operations were used in the proof of case $2b$ of the theorem about elementary transforms and solution sets?
- What property do have geometric transforms corresponding to elementary transformas if the system has no solution?