4 (Systems of) linear equations

Definition 1 (linear equation)**.** *A* linear equation *is an equation in which each term is either a constant or the product of a constant and (the first power of) a single variable. If $n \in N$, $a_1, a_2, \dots, a_n, b \in \mathbb{R}$, and $x_1, x_2, \dots, x_n$ are* unknowns *or* variables*, then a linear equation is given by*

$$\sum_{k=1}^{n} a_k x_k = b \Leftrightarrow a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b.$$

The numbers $a_1, a_2, \dots, a_n$ are called coefficients *and b is the* right-hand side*. A* solution *is an n-tuple $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ that fulfills the equation. If $b = 0$, the equation is called* homogenous *otherwise* inhomogenous*.*

Definition 2 (system of linear equations)**.** *Let $m, n \in \mathbb{N}$. A* system of linear equations *with m equations and n variables is given by*

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

where $a_{i,j}, b_i \in \mathbb{R}$ for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$. A solution *is an n-tuple $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ that fulfills all m equations of the system simultaneously. If $b_1 = b_2 = \cdots = b_m = 0$, the system is called* homogenous *otherwise* inhomogenous*.*

Problem 3. *Given: A system of linear equations.*
Goal: Find all possible solutions given as

$$S = \left\{ (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n \colon \sum_{k=1}^{n} a_{i,k} x_k = b_i \text{ for all } i = 1, 2, \dots, m \right\}.$$

Definition 4 (matrix)**.** *Let $m, n \in \mathbb{N}$. A scheme*

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ & \vdots & & \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix},$$

where $a_{i,j} \in \mathbb{R}$ for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$ is called a matrix*, more precisely an $(m \times n) -$ matrix. Furthermore, $(a_{i,1}, a_{i,2}, \dots, a_{i,n})$ are the* rows *and $(a_{1,j}, a_{2,j}, \dots, a_{m,j})^T$ are the* columns *of the matrix.*

Definition 5 ((extended) matrix of coefficients)**.** *If*

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

is a system of linear equations, we call

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ & \vdots & & \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

the matrix of coefficients *and*

$$(A|\mathbf{b}) = \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_2 \\ & \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_m \end{array}\right)$$

the extended matrix of coefficients*. We write $A\mathbf{x} = \mathbf{b}$ as an abbreviation for the system of linear equations.*

Definition 6 (simple form)**.** *An $(m \times n)$-matrix is in* simple form *if*

$$A = \begin{pmatrix} *_1 & & & & \\ & *_2 & & & \\ & & *_3 & & \\ & & & \ddots & \\ & & & & *_r \\ & & & & \end{pmatrix}$$

such that the entries at positions $*_i$ *are* $\neq 0$ *for* $i = 1, 2, \ldots, r$ *and below the line are only* 0s.

More precisely, A is in simple form if

(i) *there exists a number* $r \in \mathbb{N}$ *with* $0 \leq r \leq m$ *such that*

- *the rows with index* $1, 2, \ldots, r$ *each contain an entry* $\neq 0$*;*
- *the rows with index* $r+1, r+2, \ldots, m$ *contain only* 0s.

(ii) *Let* $j_i = \min\{j \colon a_{i,j} \neq 0\}$ *for* $1 \leq i \leq r$. *Obviously* $1 \leq j_i \leq n$ *and we require* $j_1 < j_2 < \cdots < j_r$.

Note that $r = 0$ *is possible. In this case all entries of A are* 0. *The entries* $a_{1,j_1}, a_{2,j_2}, \ldots, a_{r,j_r}$ *are called* pivots *of A.*

Proposition 7. *If A is a matrix in simple form, then* $j_i = i$ *for* $i = 1, 2, \ldots .r$ *can be obtained by rearrangement of its columns.*

Theorem 8 (solution of linear systems in simple form). *Let* $Ax = b$ *with A in simple form such that the pivots are in the first r columns, i.e.*

$$(A|b) = \left(\begin{array}{cccc|c} a_{1,1} & & & & b_1 \\ & a_{2,2} & & & b_2 \\ & & \ddots & & \vdots \\ & & & a_{r,r} & b_r \\ \hline & & & & b_{r+1} \\ & & & & \vdots \\ & & & & b_m \end{array}\right)$$

with $a_{1,1}, a_{2,2}, \ldots, a_{r,r} \neq 0$.

- *If* $b_i \neq 0$ *for an index* $r+1 \leq i \leq m$*, then* $S(A, b) = \emptyset$.
- *If* $b_i = 0$ *for* $i = r+1, r+2, \ldots, m$*, then* $S(A, b) \neq \emptyset$ *and can be computed as follows. Set* $k = n - r$*, choose* $\lambda_1, \lambda_2, \ldots, \lambda_k$ *as parameters, and set* $x_{r+1} = \lambda_1, x_{r+2} = \lambda_2, \ldots, x_n = \lambda_k$. *To compute* $x_1, x_2, \ldots, x_r$ *begin in row r:*

  $$\begin{aligned} a_{r,r}x_r + a_{r,r+1}\lambda_1 + \cdots + a_{r,n}\lambda_k &= b_r \\ \Leftrightarrow x_r &= \frac{1}{a_{r,r}}\left(b_r - a_{r,r+1}\lambda_1 - \cdots - a_{r,n}\lambda_k\right). \end{aligned}$$

  Plugging this into the $(r-1)$*st row we can compute* x_{r-1} *and so on.*

Definition 9 (free & bounded variables)**.** *In the situation of Theorem 8, the variables* $x_{r+1}, x_{r+2}, \dots, x_n$ *are called* free*, and the variables* $x_1, x_2, \dots, x_r$ *are called* bounded*.*

Definition 10 (elementary row transformations)**.** Row switching *(switching all elements of row i with its counterparts in row j),* row multiplication *(muliplying all elements in row i with a nonzero scalar), and* row addition *(adding row j multiplied with a nonzero scalar to row i) are called* elementary row transformations*.*

Theorem 11 (row transformations do not change solutions)**.** *Let* $(A|b)$ *be an augmented matrix of coefficients and suppose that* $(A'|b')$ *is obtained from* $(A|b)$ *by an elementary row transformation. Then* $S(A,b) = S(A'|b')$*, i.e.* $Ax = b$ *and* $A'x = b'$ *have the same set of solutions.*

Theorem 12 (matrices can be transformed into simple from)**.** *For every matrix* A *there is a matrix* B *in simple form that can be obtained from* A *by finitely many elementary row transformations.*

Theorem 13 (Gauss algorithm)**.** *Given a system* $Ax = b$ *of linear equations.*

1. *If* $A = 0$*,* STOP*.*

2. *Locate the first nonzero column from the left. One row has a nonzero entry in this column. If necessary, swap this line with the first.*

3. *The first nonzero entry of the first row is in the first nonzero column. Using this entry, subtract suitable multiples of the first line from the other lines to generate zero entries in that column. This yields matrix* $(A'|b')$*.*

4. *If* $(A'|b')$ *is not in simple form, repeat step 3 with* $(A|b) = (A'|b')$*. Otherwise compute* r *and check* b' *to conclude whether* $(A'|b')$ *has solutions. If yes, compute the parametrization.*

Theorem 14 (Computation of inverse matrices)**.** *The inverse of a matrix can be computed using the algorithm of* Gauss*.*

Theorem 15 (Invertible matrices & linear systems)**.** *Let* $A \in \mathbb{R}^{n\times n}$*. Then the following are equivalent.*

(i) A *is invertible.*

(ii) For every $\mathbf{b} \in \mathbb{R}^n$*:* $A\mathbf{x} = \mathbf{b}$ *has a unique solution.*

(iii) A *has rank* n*.*

Theorem 16. *Let $A \in \mathbb{R}^{m \times n}$ and $r = \operatorname{rank}(A)$. Then*

$$\dim(\operatorname{kernel}(A)) = \dim(S(A|\mathbf{0})) = n - r.$$