Chapter 6
Linear Systems of Equations

Solving linear systems of equations is a central problem of Linear Algebra that we discuss in an introductory way in this chapter. Such systems arise in numerous applications from engineering to the natural and social sciences. Major sources of linear systems of equations are the discretization of differential equations and the linearization of nonlinear equations. In this chapter we analyze the solution sets of linear systems of equations and we characterize the number of solutions using the echelon form from Chap. 5. We also develop an algorithm for the computation of the solutions.

Definition 6.1 A *linear system (of equations)* over a field K with n equations in m *unknowns* $x_1, \ldots, x_m$ has the form

$$a_{11}x_1 + \ldots + a_{1m}x_m = b_1,$$
$$a_{21}x_1 + \ldots + a_{2m}x_m = b_2,$$
$$\vdots$$
$$a_{n1}x_1 + \ldots + a_{nm}x_m = b_n$$

or

$$Ax = b,$$

where the *coefficient matrix* $A = [a_{ij}] \in K^{n,m}$ and the *right hand side* $b = [b_i] \in K^{n,1}$ are given. If $b = 0$, then the linear system is called *homogeneous*, otherwise *non-homogeneous*. Every $\widehat{x} \in K^{m,1}$ with $A\widehat{x} = b$ is called a *solution* of the linear system. All these $\widehat{x}$ form the *solution set* of the linear system, which we denote by $\mathscr{L}(A, b)$.

The next result characterizes the solution set $\mathscr{L}(A, b)$ of the linear system $Ax = b$ using the solution set $\mathscr{L}(A, 0)$ of the associated homogeneous linear system $Ax = 0$.

Lemma 6.2 *Let $A \in K^{n,m}$ and $b \in K^{n,1}$ with $\mathscr{L}(A, b) \neq \emptyset$ be given. If $\widehat{x} \in \mathscr{L}(A, b)$, then*

$$\mathscr{L}(A, b) = \widehat{x} + \mathscr{L}(A, 0) := \{\widehat{x} + \widehat{z} \mid \widehat{z} \in \mathscr{L}(A, 0)\}.$$

Proof If $\widehat{z} \in \mathscr{L}(A, 0)$, and thus $\widehat{x} + \widehat{z} \in \widehat{x} + \mathscr{L}(A, 0)$, then

$$A(\widehat{x} + \widehat{z}) = A\widehat{x} + A\widehat{z} = b + 0 = b.$$

Hence $\widehat{x} + \widehat{z} \in \mathscr{L}(A, b)$, which shows that $\widehat{x} + \mathscr{L}(A, 0) \subseteq \mathscr{L}(A, b)$.
    Let now $\widehat{x}_1 \in \mathscr{L}(A, b)$ and let $\widehat{z} := \widehat{x}_1 - \widehat{x}$. Then

$$A\widehat{z} = A\widehat{x}_1 - A\widehat{x} = b - b = 0,$$

i.e., $\widehat{z} \in \mathscr{L}(A, 0)$. Hence $\widehat{x}_1 = \widehat{x} + \widehat{z} \in \widehat{x} + \mathscr{L}(A, 0)$, which shows that $\mathscr{L}(A, b) \subseteq \widehat{x} + \mathscr{L}(A, 0)$. $\square$

We will have a closer look at the set $\mathscr{L}(A, 0)$: Clearly, $0 \in \mathscr{L}(A, 0) \neq \emptyset$. If $\widehat{z} \in \mathscr{L}(A, 0)$, then for all $\lambda \in K$ we have $A(\lambda\widehat{z}) = \lambda(A\widehat{z}) = \lambda \cdot 0 = 0$, and hence $\lambda\widehat{z} \in \mathscr{L}(A, 0)$. Furthermore, for $\widehat{z}_1, \widehat{z}_2 \in \mathscr{L}(A, 0)$ we have

$$A(\widehat{z}_1 + \widehat{z}_2) = A\widehat{z}_1 + A\widehat{z}_2 = 0 + 0 = 0,$$

and hence $\widehat{z}_1 + \widehat{z}_2 \in \mathscr{L}(A, 0)$. Thus, $\mathscr{L}(A, 0)$ is a nonempty subset of $K^{m,1}$ that is closed under scalar multiplication and addition.

Lemma 6.3 *If $A \in K^{n,m}$, $b \in K^{n,1}$ and $S \in K^{n,n}$, then $\mathscr{L}(A, b) \subseteq \mathscr{L}(SA, Sb)$. Moreover, if S is invertible, then $\mathscr{L}(A, b) = \mathscr{L}(SA, Sb)$.*

Proof If $\widehat{x} \in \mathscr{L}(A, b)$, then also $SA\widehat{x} = Sb$, and thus $\widehat{x} \in \mathscr{L}(SA, Sb)$, which shows that $\mathscr{L}(A, b) \subseteq \mathscr{L}(SA, Sb)$. If S is invertible and $\widehat{y} \in \mathscr{L}(SA, Sb)$, then $SA\widehat{y} = Sb$. Multiplying from the left with S^{-1} yields $A\widehat{y} = b$. Since $\widehat{y} \in \mathscr{L}(A, b)$, we have $\mathscr{L}(SA, Sb) \subseteq \mathscr{L}(A, b)$. $\square$

Consider the linear system of equations $Ax = b$. By Theorem 5.2 we can find a matrix $S \in GL_n(K)$ such that SA is in echelon form. Let $\widetilde{b} = [\widetilde{b}_i] := Sb$, then $\mathscr{L}(A, b) = \mathscr{L}(SA, \widetilde{b})$ by Lemma 6.3, and the linear system $SAx = \widetilde{b}$ takes the form

Suppose that $\text{rank}(A) = r$, and let $j_1, j_2, \ldots, j_r$ be the pivot columns. Using a right-multiplication with a permutation matrix we can move the r pivot columns of SA to the first r columns. This is achieved by

$$P^T := [e_{j_1}, \ldots, e_{j_r}, e_1, \ldots, e_{j_1-1}, e_{j_1+1}, \ldots, e_{j_2-1}, e_{j_2+1}, \ldots, e_{j_r-1}, e_{j_r+1}, \ldots, e_m] \in K^{m,m},$$

which yields

$$\widetilde{A} := SAP^T = \begin{bmatrix} I_r & \widetilde{A}_{12} \\ 0_{n-r,r} & 0_{n-r,m-r} \end{bmatrix},$$

where $\widetilde{A}_{12} \in K^{r,m-r}$. If $r = m$, then we have $\widetilde{A}_{12} = [\ ]$. This permutation leads to a simplification of the following presentation, but it is usually omitted in practical computations.

Since $P^T P = I_m$, we can write $SAx = \widetilde{b}$ as $(SAP^T)(Px) = \widetilde{b}$, or $\widetilde{A}y = \widetilde{b}$, which has the form

$$\underbrace{\begin{bmatrix} I_r & \widetilde{A}_{12} \\ \hline 0_{n-r,r} & 0_{n-r,m-r} \end{bmatrix}}_{=\widetilde{A}:=SAP^T} \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_r \\ y_{r+1} \\ \vdots \\ y_m \end{bmatrix}}_{=y:=Px} = \underbrace{\begin{bmatrix} \widetilde{b}_1 \\ \vdots \\ \widetilde{b}_r \\ \widetilde{b}_{r+1} \\ \vdots \\ \widetilde{b}_n \end{bmatrix}}_{=\widetilde{b}:=Sb}. \tag{6.1}$$

The left-multiplication of x with P just means a different ordering of the unknowns $x_1, \ldots, x_m$. Thus, the solutions of $Ax = b$ can be easily recovered from the solutions of $\widetilde{A}y = \widetilde{b}$, and vice versa: We have $\widehat{y} \in \mathcal{L}(\widetilde{A}, \widetilde{b})$ if and only if $\widehat{x} := P^T\widehat{y} \in \mathcal{L}(SA, \widetilde{b}) = \mathcal{L}(A, b)$.

The solutions of (6.1) can now be determined using the *extended coefficient matrix*

$$[\widetilde{A}, \widetilde{b}] \in K^{n,m+1},$$

which is obtained by attaching $\widetilde{b}$ as an extra column to $\widetilde{A}$. Note that $\text{rank}(\widetilde{A}) \leq \text{rank}([\widetilde{A}, \widetilde{b}])$, with equality if and only if $\widetilde{b}_{r+1} = \cdots = \widetilde{b}_n = 0$.

If $\text{rank}(\widetilde{A}) < \text{rank}([\widetilde{A}, \widetilde{b}])$, then at least one of $\widetilde{b}_{r+1}, \ldots, \widetilde{b}_n$ is nonzero. Then (6.1) cannot have a solution, since all entries of $\widetilde{A}$ in the rows $r + 1, \ldots, n$ are zero.

If, on the other hand, $\text{rank}(\widetilde{A}) = \text{rank}([\widetilde{A}, \widetilde{b}])$, then $\widetilde{b}_{r+1} = \cdots = \widetilde{b}_n = 0$ and (6.1) can be written as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix} = \begin{bmatrix} \widetilde{b}_1 \\ \vdots \\ \widetilde{b}_r \end{bmatrix} - \widetilde{A}_{12} \begin{bmatrix} y_{r+1} \\ \vdots \\ y_m \end{bmatrix}. \tag{6.2}$$

This representation implies, in particular, that

$$\widetilde{b}^{(m)} := [\widetilde{b}_1, \ldots, \widetilde{b}_r, \underbrace{0, \ldots, 0}_{m-r}]^T \in \mathscr{L}(\widetilde{A}, \widetilde{b}) \neq \emptyset.$$

From Lemma 6.2 we know that $\mathscr{L}(\widetilde{A}, \widetilde{b}) = \widetilde{b}^{(m)} + \mathscr{L}(\widetilde{A}, 0)$. In order to determine $\mathscr{L}(\widetilde{A}, 0)$ we set $\widetilde{b}_1 = \cdots = \widetilde{b}_r = 0$ in (6.2), which yields

$$\mathscr{L}(\widetilde{A}, 0) = \left\{ [\widehat{y}_1, \ldots, \widehat{y}_m]^T \mid \widehat{y}_{r+1}, \ldots, \widehat{y}_m \text{ arbitrary and} \right. \tag{6.3}$$
$$\left. [\widehat{y}_1, \ldots, \widehat{y}_r]^T = 0 - \widetilde{A}_{12}[\widehat{y}_{r+1}, \ldots, \widehat{y}_m]^T \right\}.$$

If $r = m$, then $\widetilde{A}_{12} = [\ ]$, $\mathscr{L}(\widetilde{A}, 0) = \{0\}$ and thus $|\mathscr{L}(\widetilde{A}, \widetilde{b})| = 1$, i.e., the solution of $\widetilde{A}y = \widetilde{b}$ is uniquely determined.

Example 6.4 For the extended coefficient matrix

$$[\widetilde{A}, \widetilde{b}] = \begin{bmatrix} 1 & 0 & 3 & \widetilde{b}_1 \\ 0 & 1 & 4 & \widetilde{b}_2 \\ 0 & 0 & 0 & \widetilde{b}_3 \end{bmatrix} \in \mathbb{Q}^{3,4}$$

we have $\text{rank}(\widetilde{A}) = \text{rank}([\widetilde{A}, \widetilde{b}])$ if and only if $\widetilde{b}_3 = 0$. If $\widetilde{b}_3 \neq 0$, then $\mathscr{L}(\widetilde{A}, \widetilde{b}) = \emptyset$.

If $\widetilde{b}_3 = 0$, then $\widetilde{A}y = \widetilde{b}$ can be written as

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \widetilde{b}_1 \\ \widetilde{b}_2 \end{bmatrix} - \begin{bmatrix} 3 \\ 4 \end{bmatrix} [y_3].$$

Hence, $\widetilde{b}^{(3)} = [\widetilde{b}_1, \widetilde{b}_2, 0]^T \in \mathscr{L}(\widetilde{A}, \widetilde{b})$ and

$$\mathscr{L}(\widetilde{A}, 0) = \left\{ [\widehat{y}_1, \widehat{y}_2, \widehat{y}_3]^T \mid \widehat{y}_3 \text{ arbitrary and } [\widehat{y}_1, \widehat{y}_2]^T = -[3, 4]^T[\widehat{y}_3] \right\}.$$

Summarizing our considerations we have the following algorithm for solving a linear system of equations.

Algorithm 6.5 Let $A \in K^{n,m}$ and $b \in K^{n,1}$ be given.

(1) Determine $S \in GL_n(K)$ such that SA is in echelon form and define $\widetilde{b} := Sb$.
(2a) If $\text{rank}(SA) < \text{rank}([SA, \widetilde{b}])$, then $\mathscr{L}(SA, \widetilde{b}) = \mathscr{L}(A, b) = \emptyset$.
(2b) If $r = \text{rank}(A) = \text{rank}([SA, \widetilde{b}])$, then define $\widetilde{A} := SAP^T$ as in (6.1).

We have $\widetilde{b}^{(m)} \in \mathscr{L}(\widetilde{A}, \widetilde{b})$ and $\mathscr{L}(\widetilde{A}, \widetilde{b}) = \widetilde{b}^{(m)} + \mathscr{L}(\widetilde{A}, 0)$, where $\mathscr{L}(\widetilde{A}, 0)$ is determined as in (6.3), as well as $\mathscr{L}(A, b) = \{P^T \widehat{y} \mid \widehat{y} \in \mathscr{L}(\widetilde{A}, \widetilde{b})\}$.

Since $\text{rank}(A) = \text{rank}(SA) = \text{rank}(\widetilde{A})$ and $\text{rank}([A, b]) = \text{rank}([SA, \widetilde{b}]) = \text{rank}([\widetilde{A}, \widetilde{b}])$, the discussion above also yields the following result about the different cases of the solvability of a linear system of equations.

Corollary 6.6 *For $A \in K^{n,m}$ and $b \in K^{n,1}$ the following assertions hold:*

(1) If $\operatorname{rank}(A) < \operatorname{rank}([A, b])$, *then* $\mathscr{L}(A, b) = \emptyset$.
(2) If $\operatorname{rank}(A) = \operatorname{rank}([A, b]) = m$, *then* $|\mathscr{L}(A, b)| = 1$ *(i.e., there exists a unique solution).*
(3) If $\operatorname{rank}(A) = \operatorname{rank}([A, b]) < m$, *then there exist many solutions. If the field K has infinitely many elements (e.g., when $K = \mathbb{Q}$, $K = \mathbb{R}$ or $K = \mathbb{C}$), then there exist infinitely many pairwise distinct solutions.*

The different cases in Corollary 6.6 will be studied again in Example 10.8.

Example 6.7 Let $K = \mathbb{Q}$ and consider the linear system of equations $Ax = b$ with

$$
A = \begin{bmatrix} 1\ 2\ 2\ 1 \\ 0\ 1\ 0\ 3 \\ 1\ 0\ 3\ 0 \\ 2\ 3\ 5\ 4 \\ 1\ 1\ 3\ 3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \\ 2 \end{bmatrix}.
$$

We form $[A, b]$ and apply the Gaussian elimination algorithm in order to transform A into echelon form:

$$
[A, b] \rightsquigarrow \left[\begin{array}{rrrr|r} 1 & 2\ 2 & 1 & 1 \\ 0 & 1\ 0 & 3 & 0 \\ 0 & -2\ 1 & -1 & 1 \\ 0 & -1\ 1 & 2 & 1 \\ 0 & -1\ 1 & 2 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{rrrr|r} 1\ 2\ 2\ 1 & 1 \\ 0\ 1\ 0\ 3 & 0 \\ 0\ 0\ 1\ 5 & 1 \\ 0\ 0\ 1\ 5 & 1 \\ 0\ 0\ 1\ 5 & 1 \end{array}\right]
$$

$$
\rightsquigarrow \left[\begin{array}{rrrr|r} 1\ 2\ 2\ 1 & 1 \\ 0\ 1\ 0\ 3 & 0 \\ 0\ 0\ 1\ 5 & 1 \\ 0\ 0\ 0\ 0 & 0 \\ 0\ 0\ 0\ 0 & 0 \end{array}\right] \rightsquigarrow \left[\begin{array}{rrrr|r} 1\ 0\ 2 & -5 & 1 \\ 0\ 1\ 0 & 3 & 0 \\ 0\ 0\ 1 & 5 & 1 \\ 0\ 0\ 0 & 0 & 0 \\ 0\ 0\ 0 & 0 & 0 \end{array}\right]
$$

$$
\rightsquigarrow \left[\begin{array}{rrrr|r} 1\ 0\ 0 & -15 & -1 \\ 0\ 1\ 0 & 3 & 0 \\ 0\ 0\ 1 & 5 & 1 \\ 0\ 0\ 0 & 0 & 0 \\ 0\ 0\ 0 & 0 & 0 \end{array}\right] = [SA|\tilde{b}].
$$

Here $\operatorname{rank}(SA) = \operatorname{rank}([SA, \tilde{b}]) = 3$, and hence there exist solutions. The pivot columns are $j_i = i$ for $i = 1, 2, 3$, so that $P = P^T = I_4$ and $\tilde{A} = SA$. Now $SAx = \tilde{b}$ can be written as

$$
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} -15 \\ 3 \\ 5 \end{bmatrix} [x_4].
$$

Consequently, $\widetilde{b}^{(4)} = [-1, 0, 1, 0]^T \in \mathscr{L}(A, b)$ and $\mathscr{L}(A, b) = \widetilde{b}^{(4)} + \mathscr{L}(A, 0)$, where

$$\mathscr{L}(A, 0) = \left\{ [\widehat{x}_1, \ldots, \widehat{x}_4]^T \mid \widehat{x}_4 \text{ arbitrary and } [\widehat{x}_1, \widehat{x}_2, \widehat{x}_3]^T = -[-15, 3, 5]^T [\widehat{x}_4] \right\}.$$

Exercises

6.1 Find a field K and matrices $A \in K^{n,m}$, $S \in K^{n,n}$ and $b \in K^{n,1}$ with $\mathscr{L}(A, b) \neq \mathscr{L}(SA, Sb)$.

6.2 Determine $\mathscr{L}(A, b)$ for the following A and b:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 6 \end{bmatrix} \in \mathbb{R}^{3,3}, \quad b = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \in \mathbb{R}^{3,1},$$

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & -1 & -1 \\ 1 & -1 & 6 & 2 \end{bmatrix} \in \mathbb{R}^{3,4}, \quad b = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \in \mathbb{R}^{3,1},$$

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 6 \\ 1 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{4,3}, \quad b = \begin{bmatrix} 1 \\ -2 \\ 3 \\ 1 \end{bmatrix} \in \mathbb{R}^{4,1},$$

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 6 \\ 1 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{4,3}, \quad b = \begin{bmatrix} 1 \\ -2 \\ 3 \\ 0 \end{bmatrix} \in \mathbb{R}^{4,1}.$$

6.3 Let $\alpha \in \mathbb{Q}$,

$$A = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \in \mathbb{Q}^{3,3}, \quad b_\alpha = \begin{bmatrix} 6 \\ 3 \\ \alpha \end{bmatrix} \in \mathbb{Q}^{3,1}.$$

Determine $\mathscr{L}(A, 0)$ and $\mathscr{L}(A, b_\alpha)$ in dependence of α.

6.4 Let $A \in K^{n,m}$ and $B \in K^{n,s}$. For $i = 1, \ldots, s$ denote by b_i the ith column of B. Show that the linear system of equations $AX = B$ has at least one solution $\widehat{X} \in K^{m,s}$ if and only if

$$\text{rank}(A) = \text{rank}([A, b_1]) = \text{rank}([A, b_2]) = \cdots = \text{rank}([A, b_s]).$$

Find conditions under which this solution is unique.

6.5   Let

$$A = \begin{bmatrix} 0 & \beta_1 & & \\ \alpha_2 & 0 & \ddots & \\ & \ddots & \ddots & \beta_n \\ & & \alpha_n & 0 \end{bmatrix} \in K^{n,n}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \in K^{n,1}$$

be given with β_i, $\alpha_i \neq 0$ for all i. Determine a recursive formula for the entries of the solution of the linear system $Ax = b$.