Note 6

Systems of linear equations

6.1 Structure of systems of linear equations

When dealing with an equation in one variable, it is very common to use the variable x. In Example 1.17, we studied for example the equation $2|x| = 2x + 1$. Often there is not just one variable, but several. If there are two variables, one often uses x and y, if there are three x, y and z, but what to do if there are more variables, say five variables? In such cases it is common to use variables x_1, x_2, etcetera. For example, if we need five variables, we just use x_1, x_2, x_3, x_4 and x_5. We can even leave the precise number of variables unspecified and say that we have n variables for some natural number $n \in \mathbb{N}$. One says that one has an equation in the n variables $x_1, \ldots, x_n$.

A *linear equation* in the n variables $x_1, \ldots, x_n$ is an equation of the form

$$a_1 \cdot x_1 + \cdots a_n \cdot x_n = b,$$

where $a_1, \ldots, a_n, b$ are constants. These constants will typically be real or complex numbers, depending on the situation. To avoid having to specify all the time if we are working with real or complex numbers, let us introduce the following definition:

Definition 6.1

A set $\mathbb{F}$ is called a *field*, if there is an addition $+$ and multiplication $\cdot$ defined for all pairs of elements of $\mathbb{F}$ in such a way that the following rules are satisfied:

1. Addition and multiplication are *associative*: $a_1 + (a_2 + a_3) = (a_1 + a_2) + a_3$, and $a_1 \cdot (a_2 \cdot a_3) = (a_1 \cdot a_2) \cdot a_3$ for all $a_1, a_2, a_3 \in \mathbb{F}$.

2. Addition and multiplication are *commutative*: $a_1 + a_2 = a_2 + a_1$, and $a_1 \cdot a_2 = a_2 \cdot a_1$ for all $a_1, a_2, a_3 \in \mathbb{F}$.

3. *Distributivity* of multiplication over addition holds: $a_1 \cdot (a_2 + a_3) = a_1 \cdot a_2 + a_1 \cdot a_3$ for all $a_1, a_2, a_3 \in \mathbb{F}$.

4. Addition and multiplication have a neutral element, that is to say certain elements in $\mathbb{F}$ usually denoted by 0 and 1 that satisfy $a + 0 = a$ and $a \cdot 1 = a$ for all $a \in \mathbb{F}$.

5. Additive inverses exist: for every $a \in \mathbb{F}$, there exists an element in $\mathbb{F}$, denoted by $-a$ and called the additive inverse of a, such that $a + (-a) = 0$.

6. Multiplicative inverses exist: for every $a \in \mathbb{F} \setminus \{0\}$, there exists an element in $\mathbb{F}$, denoted by a^{-1} or $1/a$ and called the multiplicative inverse of a, such that $a \cdot a^{-1} = 1$.

Theorems 3.10 and 3.11 together simply state that the complex numbers form a field. Also the real numbers $\mathbb{R}$ with the usual addition and multiplication form a field. There are many more possible examples of fields, but whenever we use the symbol $\mathbb{F}$ or write something like "the field $\mathbb{F}$", you can just think of $\mathbb{R}$ or $\mathbb{C}$. Just to show that there exist more fields, we give two examples.

Example 6.2

Let $\mathbb{F} = \mathbb{Q}$ be the set of rational numbers, see Example 2.4. This set, equipped with the usual addition and multiplication, is a field. It is called the field of rational numbers.

|||| Example 6.3

Let $\mathbb{F}_2 = \{0,1\}$ and define addition and multiplication as follows: $0+0=0, 0+1=1, 1+0=1, 1+1=0$ and $0\cdot 0=0, 0\cdot 1=0, 1\cdot 0=0, 1\cdot 1=1$. Then with this addition and multiplication, $\mathbb{F}_2$ is a field. It is called the field of bits, the binary field, or also the finite field with two elements.

Returning to our study of linear equations, we can now give a more precise definition.

|||| Definition 6.4

A linear equation over a field $\mathbb{F}$ in the n variables $x_1, \dots, x_n$, is an equation of the form

$$a_1 \cdot x_1 + \cdots + a_n \cdot x_n = b,$$

where $a_1, \dots, a_n, b \in \mathbb{F}$.
A solution to this linear equation is an n-tuple $(v_1, \dots, v_n) \in \mathbb{F}^n$ such that $a_1 \cdot v_1 + \cdots + a_n \cdot v_n = b$.

We have seen the notation $\mathbb{F}^n$ in this definition before in Section 2.1, see equation (2-3). It is the Cartesian product of $\mathbb{F}$ with itself n times. More down to earth, $\mathbb{F}^n$ is simply the set of all n-tuples $(v_1, \dots, v_n)$, where each coordinate is an element from $\mathbb{F}$. Sometimes the multiplication between the constant and variables are omitted. For example $2x_1$ has the same meaning as $2 \cdot x_1$.

There is a subtlety in Definition 6.4 that is easy to miss. If we say that we consider a linear equation over $\mathbb{F}$, we are only interested in solutions $(v_1, \dots, v_n)$ that lie in $\mathbb{F}^n$. In other words, by specifying that the linear equation is over $\mathbb{F}$, we implicitly say that all the coordinates of a solution $(v_1, \dots, v_n)$ must lie in $\mathbb{F}$. Let us consider a few examples.

|||| Example 6.5

1. Find a solution to the linear equation $3x_1 + x_2 = 5$ over $\mathbb{R}$.
2. Consider the linear equation $x_1 + x_2 = 0$ over $\mathbb{C}$. Is $(i, -i) \in \mathbb{C}^2$ a solution to this linear equation?
3. Consider the linear equation $x_1 + x_2 = 0$ over $\mathbb{R}$. Is $(i, -i) \in \mathbb{C}^2$ a solution to this linear equation?

4. Consider the linear equation $x_1 + x_2 = 0$ over $\mathbb{R}$. Find a solution.

Answer:

1. There are many possible solutions, but for example $(v_1, v_2) = (0, 5)$ is a solution, since $3 \cdot 0 + 5 = 5$.
2. Since $i + (-i) = 0$, the pair $(i, -i) \in \mathbb{C}^2$ is indeed a solution to the linear equation $x_1 + x_2 = 0$ over $\mathbb{C}$.
3. Even though $i + (-i) = 0$, the pair $(i, -i)$ is not a solution to the linear equation $x_1 + x_2 = 0$ over $\mathbb{R}$. The reason is that the pair $(-i, i)$ is not an element of $\mathbb{R}^2$.
4. A possible solution is $(1, -1)$. Another solution is $(0, 0)$.

Now we arrive at the main topic of this section, namely systems of linear equations. It is simply an extension of Definition 6.4 by not considering only one linear equation, but several linear equations over a field $\mathbb{F}$ at the same time.

Definition 6.6

A system of m linear equations $R_1, \ldots, R_m$ over a field $\mathbb{F}$ in the n variables $x_1, \ldots, x_n$, is a system of m equations of the form

$$\begin{cases} \begin{array}{rccccccc} R_1: & a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & b_1 \\ R_2: & a_{21} \cdot x_1 & + & \cdots & + & a_{2n} \cdot x_n & = & b_2 \\ & & & \vdots & & & & \vdots \\ R_m: & a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & b_m \end{array} \end{cases}$$

where $a_{11}, \ldots, a_{mn}, b_1, \ldots, b_m \in \mathbb{F}$.
A solution to this system of linear equations is an n-tuple $(v_1, \ldots, v_n) \in \mathbb{F}^n$ such that for all j between 1 and m it holds that $a_{j1} \cdot v_1 + \cdots + a_{jn} \cdot v_n = b_j$.

Some explanation of the notation is in order. First of all, a double index was used for the constants in front of the variables. The constant a_{ij} denotes the constant occurring in equation i in front of the variable x_j. For example, if we have at least two equations and at least three variables, then a_{23} would denote the constant in the second equation in front of the variable x_3. In case $m = 1$ in Definition 6.6, we just recover the case of one linear equation as described in Definition 6.4.

The use of the brace $\{$ in front of the equations is just to emphasize that all equations are considered simultaneously and that a solution to the system should satisfy all equations at the same time. In logical terms, we can therefore write that an n-tuple $(v_1, \ldots, v_n) \in \mathbb{F}^n$ is a solution to the system of equations as given in Definition 6.6 precisely if:

$$a_{11} \cdot v_1 + \cdots + a_{1n} \cdot v_n = b_1 \quad \wedge \quad \cdots \quad \wedge \quad a_{m1} \cdot v_1 + \cdots + a_{mn} \cdot v_n = b_m.$$

Using $R_1, \ldots, R_m$ as "labels" for the equations, is not necessary and often these labels are just omitted. We will usually also omit these labels, but when developing the theory on how to solve systems of linear equations, they can be quite convenient. To digest this definition, let us immediately consider some examples.

|||| Example 6.7

Determine the set of solutions to the following system of two linear equations in two variables over $\mathbb{R}$:

$$\begin{cases} x_1 + 2x_2 & = & 1 \\ x_2 & = & 2 \end{cases}$$

This system is quite simple to solve, since the second equation already determines x_2 (namely $x_2 = 2$). Then using this in the first equation, we see that any pair (x_1, x_2) that satisfies *both* linear equations, will satisfy $x_2 = 2$ and $x_1 = 1 - 2x_2 = 1 - 2 \cdot 2 = -3$. Hence the system has only one solution, namely $(x_1, x_2) = (-3, 2)$. The set of all solutions is therefore given by $\{(-3, 2)\}$.

|||| Example 6.8

Consider the following system of linear equations over $\mathbb{R}$ in the variables $x_1, \ldots, x_4$:

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

Let us see how this example fits with Definition 6.6. First of all, we have two linear equations and hence $m = 2$. Further, the only variables occurring in these three equations are x_1, x_2, x_3 and x_4. Hence we can choose $n = 4$. To determine the a_{ij} is now a matter of reading off the constants in front of the variables. However, before we do this, it is convenient to rewrite the system of equations a bit as follows:

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

We can now read off directly that $a_{11} = 2$, $a_{12} = 5$, $a_{13} = 0$, $a_{14} = 1$, $b_1 = 0$, $a_{21} = 3$, $a_{22} = 0$, $a_{23} = -1$, $a_{24} = 0$ and $b_2 = 6$. We will determine the solutions of this system of linear equations later.

A system of m linear equations over a field $\mathbb{F}$ in the n variables $x_1, \ldots, x_n$ is called *homogeneous*, if for all i between 1 and m, it holds that $b_i = 0$. Otherwise, the system is called *inhomogeneous*. The system given in Example 6.8 is inhomogeneous, since in that example $b_2 \neq 0$. An example of a homogeneous system of linear equations in three variables is:

$$\begin{cases} 3 \cdot x_1 & + & 5 \cdot x_2 & + & 10 \cdot x_3 & = & 0 \\ 5 \cdot x_1 & + & 2 \cdot x_2 & - & 2 \cdot x_3 & = & 0 \end{cases}$$

Note that the all-zero tuple $(0, 0, 0)$ is a possible solution to this system. More generally, one can show that a homogeneous system of linear equations in n variables has the all-zero n-tuple $(0, \ldots, 0)$ as solution. Let us end this section by giving two structure theorems concerning the solution sets of systems of linear equations. One will be for homogeneous systems, one for inhomogeneous systems.

Theorem 6.9

Let a homogeneous system of m linear equations $R_1, \ldots, R_m$ over a field $\mathbb{F}$ in the n variables $x_1, \ldots, x_n$ be given, say

$$\begin{cases} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & 0 \\ a_{21} \cdot x_1 & + & \cdots & + & a_{2n} \cdot x_n & = & 0 \\ & & \vdots & & & & \vdots \\ a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & 0 \end{cases}$$

where $a_{11}, \ldots, a_{mn} \in \mathbb{F}$. Then

1. The all-zero tuple $(0, \ldots, 0) \in \mathbb{F}^n$ is a solution to the system.
2. If $(v_1, \ldots, v_n) \in \mathbb{F}^n$ is a solution and $c \in \mathbb{F}$, then $(c \cdot v_1, \ldots, c \cdot v_n)$ is also a solution.
3. If $(v_1, \ldots, v_n), (w_1, \ldots, w_n) \in \mathbb{F}^n$ are solutions, then $(v_1 + w_1, \ldots, v_n + w_n)$ is also a solution.

Proof. We have already remarked that the all-zero tuple is a solution to a homogeneous system. We will prove the third statement and leave proving the second statement to the reader. If $(v_1, \ldots, v_n), (w_1, \ldots, w_n) \in \mathbb{F}^n$ are solutions, then we know that for all j between 1 and m that:

$$a_{j1} \cdot v_1 + \cdots + a_{jn} \cdot v_n = 0 \text{ and } a_{j1} \cdot w_1 + \cdots + a_{jn} \cdot w_n = 0.$$

Adding these equations, we find that

$$a_{j1} \cdot v_1 + a_{j1} \cdot w_1 + \cdots + a_{jn} \cdot v_n + a_{jn} \cdot w_n = 0,$$

which can be rewritten as

$$a_{j1} \cdot (v_1 + w_1) + \cdots + a_{jn} \cdot (v_n + w_n) = 0.$$

The reader is encouraged to think about which properties of a field from Definition 6.1 we have used here. Since this is true for any j, we may conclude that $(v_1 + w_1, \ldots, v_n + w_n)$ is also a solution to the given homogeneous system of linear equations. □

Theorem 6.10

Let an inhomogeneous system of m linear equations $R_1, \ldots, R_m$ over a field $\mathbb{F}$ in the n variables $x_1, \ldots, x_n$ be given, say

$$\left\{\begin{array}{ccccccc} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & b_1 \\ a_{21} \cdot x_1 & + & \cdots & + & a_{2n} \cdot x_n & = & b_2 \\ & & \vdots & & & & \vdots \\ a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & b_m \end{array}\right.$$

where $a_{11}, \ldots, a_{mn}, b_1, \ldots, b_m \in \mathbb{F}$ and not all b_i are zero. If the system does have a solution, say $(v_1, \ldots, v_n) \in \mathbb{F}^n$, then any other solution is of the form $(v_1 + w_1, \ldots, v_n + w_n)$, where $(w_1, \ldots, w_n) \in \mathbb{F}^n$ is a solution to the corresponding homogeneous system:

$$\left\{\begin{array}{ccccccc} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & 0 \\ a_{21} \cdot x_1 & + & \cdots & + & a_{2n} \cdot x_n & = & 0 \\ & & \vdots & & & & \vdots \\ a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & 0 \end{array}\right.$$

Proof. Suppose that the system has a solution, say $(v_1, \ldots, v_n) \in \mathbb{F}^n$. Let $(v'_1, \ldots, v'_n) \in \mathbb{F}^n$ be any other solution. First of all, if we define $w_i = v'_i - v_i$, then by definition of the w_i, we obtain that $(v'_1, \ldots, v'_n) = (v_1 + w_1, \ldots, v_n + w_n)$. Hence what we need to show is that the tuple $(w_1, \ldots, w_n)$ is a solution to the homogeneous system stated in the theorem. However, we know that for all j:

$$a_{j1} \cdot v'_1 + \cdots + a_{jn} \cdot v'_n = b_j \text{ and } a_{j1} \cdot v_1 + \cdots + a_{jn} \cdot v_n = b_j.$$

Taking the difference of these two equations, we obtain:

$$a_{j1} \cdot v'_1 - a_{j1} \cdot v_1 + \cdots + a_{jn} \cdot v'_n - a_{jn} \cdot v_n = b_j - b_j,$$

which can be rewritten as

$$a_{j1} \cdot (v'_1 - v_1) + \cdots + a_{jn} \cdot (v'_n - v_n) = 0.$$

Since $w_i = v'_i - v_i$, we obtain that for all j it holds that

$$a_{j1} \cdot w_1 + \cdots + a_{jn} \cdot w_n = 0.$$

This is exactly the same as saying that $(w_1, \ldots, w_n)$ is a solution to the homogeneous system given in the theorem. $\square$

It is not a coincidence that Theorem 6.10 is formulated as it is. Indeed, the theorem holds if there exists a solution to the inhomogeneous system, but there is no guarantee that such a solution actually exists. A solution to an inhomogeneous system of linear equations is sometimes called a *particular solution*. Theorem 6.10 can then in words be described as stating that all solutions to an inhomogeneous system can be obtained as the sum of a given particular solution (if it exists) and the solutions to the corresponding homogeneous system.

Let us for the sake of completeness, give a small example of an inhomogeneous system of linear equations that has no solutions:

|||| Example 6.11

Consider the following system of two linear equations in two variables over $\mathbb{R}$:

$$\begin{cases} x_1 & + & x_2 & = & 1 \\ x_1 & + & x_2 & = & 0 \end{cases}.$$

This system is inhomogeneous, since the right-hand side of the first equation is not a zero. This system has no solutions, since it is not possible that $x_1 + x_2$ is equal to 1 and 0 at the same time! Indeed if that would be possible, we could conclude that $0 = 1$, which would be a contradiction.

To make Theorems 6.9 and 6.10 constructive, we need to figure out a way to answer the following three questions:

1. How do we describe all solutions to a homogeneous system of linear equations explicitly?
2. How do we decide if an inhomogeneous system of linear equations has a solution?
3. If it exists, how do we explicitly find a solution to an inhomogeneous system of linear equations?

Note that if we can answer these questions, Theorem 6.10 can be used to describe all solutions to an inhomogeneous system of linear equations that have at least one solution. In the next sections we will answer these questions.

6.2 Transforming a system of linear equations

In this section, we will come up with a procedure that transforms a given system of linear equations into a simpler one, without changing the solutions they have. In other words, we want to find a way to replace a possibly complicated looking system of linear equations with another, much simpler system of linear equations, but we want that the initial, possibly complicated, system has exactly the same solutions as the simpler one.

Before we start with that though, we will introduce a compact way to describe a system of linear equations using what are known as *matrices*. For now you can think of a matrix as a rectangular scheme containing elements from the field $\mathbb{F}$ one is working over. In a later chapter, we will have a more in depth discussion of matrices.

Definition 6.12

Given a linear system

$$\begin{cases} a_{11} \cdot x_1 & + & \cdots & + & a_{1n} \cdot x_n & = & b_1 \\ a_{21} \cdot x_1 & + & \cdots & + & a_{2n} \cdot x_n & = & b_2 \\ & & \vdots & & & & \vdots \\ a_{m1} \cdot x_1 & + & \cdots & + & a_{mn} \cdot x_n & = & b_m \end{cases}$$

we denote by

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

the *coefficient matrix* of the system of linear equations. The matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{bmatrix}$$

is called the *augmented matrix* of the system of linear equations.

Example 6.13

Let us consider the system of linear equations as given in Example 6.8. The coefficient matrix of this system is given by

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

while the augmented matrix of this system is

$$\begin{bmatrix} 2 & 5 & 0 & 1 & 0 \\ 3 & 0 & -1 & 0 & 6 \end{bmatrix}.$$

Sometimes one writes

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

for the augmented matrix to emphasize that the final 0 and 6 come from the right-hand side of the system of linear equations. This is just an esthetic choice though.

One says that a matrix has *rows* and *columns*. A row is a horizontal slice of a matrix, a column a vertical slice. For example, the matrix

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

has two rows: the first row is given by $[2\ 6\ 0\ 1\ 0]$, while the second row is given by $[4\ 0\ -1\ 0\ 6]$. Similarly, it has five columns, namely

$$\begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \text{ and } \begin{bmatrix} 0 \\ 6 \end{bmatrix}.$$

A matrix is said to be an $m \times n$ matrix, if it has precisely m rows and precisely n columns. Hence the matrix we just considered is a 2×5 matrix. If we consider the matrices in Definition 6.12, we see that the coefficient matrix of a system of m linear equations in n variables is an $m \times n$ matrix. Similarly, its augmented matrix is an $m \times (n+1)$ matrix. Indeed, it has one more column than the coefficient matrix, containing the b_i from the right-hand sides of the linear equations.

Now let us return to our goal: transforming a system of linear equations over a field $\mathbb{F}$ into a simpler one, without changing the solution set. The idea is to gradually transform any given system over $\mathbb{F}$ into a much simpler system, at each step making sure that the set of solutions did not change. The operations that we will use to transform the systems will consist of three types:

1. Interchange two equations.
2. Multiply a given equation with a nonzero constant from $\mathbb{F}$.
3. Add a multiple of one equation to another.

Let us explain, what these three operations do in more detail. The first one takes two linear equations from a given system, say R_i and R_j, and interchanges them. This means that after the operation R_j occurs in position i and R_i in position j. We denote this operation by $R_i \leftrightarrow R_j$.

|||| Example 6.14

Let us illustrate the interchange operation on the system from Example 6.8:

$$\begin{cases} 2 \cdot x_1 + 6 \cdot x_2 + 0 \cdot x_3 + 1 \cdot x_4 = 0 \\ 4 \cdot x_1 + 0 \cdot x_2 + (-1) \cdot x_3 + 0 \cdot x_4 = 6 \end{cases}.$$

In this case, we can perform the operation $R_1 \leftrightarrow R_2$ and obtain the system

$$\left\{ \begin{array}{rcrcrcrcr} 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \\ 2 \cdot x_1 & + & 6 \cdot x_2 & + & 0 \cdot x_3 & + & 1 \cdot x_4 & = & 0 \end{array} \right. .$$

If we, which in fact is more convenient, work with the augmented matrix of this system, the effect of the operation $R_1 \leftrightarrow R_2$ is that the augmented matrix

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

is replaced by

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

Hence the operation $R_1 \leftrightarrow R_2$ simply interchanges the first and the second row of the augmented matrix. We will usually write this as follows:

$$\begin{bmatrix} 2 & 6 & 0 & 1 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix} \underset{R_1 \leftrightarrow R_2}{\longrightarrow} \begin{bmatrix} 4 & 0 & -1 & 0 & 6 \\ 2 & 6 & 0 & 1 & 0 \end{bmatrix}.$$

The second operation we will use to simplify systems just multiplies one of the given linear equations with a *nonzero* constant $c \in \mathbb{F}$ (in other words $c \in \mathbb{F} \setminus \{0\}$). This simply means that one replaces the linear equation R_j, say given by $a_{j1}x_1 + \cdots + a_{jn}x_n = b_j$, with the linear equation $ca_{j1}x_1 + \cdots + ca_{jn}x_n = cb_j$ (which is for simplicity just denoted by $c \cdot R_j$). We denote this operation by $R_j \leftarrow c \cdot R_j$.

|||| Example 6.15

Let us illustrate the operation $R_1 \leftarrow (1/2) \cdot R_1$ on the system from Example 6.8. This amounts to replacing the system

$$\left\{ \begin{array}{rcrcrcrcr} 2 \cdot x_1 & + & 6 \cdot x_2 & + & 0 \cdot x_3 & + & 1 \cdot x_4 & = & 0 \\ 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{array} \right.$$

by

$$\left\{ \begin{array}{rcrcrcrcr} 1 \cdot x_1 & + & 3 \cdot x_2 & + & 0 \cdot x_3 & + & 1/2 \cdot x_4 & = & 0 \\ 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{array} \right. .$$

In matrix notation, we obtain

$$\begin{bmatrix} 2 & 6 & 0 & 1 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix} \underset{R_1 \leftarrow (1/2) \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix}.$$

Hence the effect of the operation $R_1 \leftarrow (1/2) \cdot R_1$ on the augmented matrix is that all entries in the first row are multiplied with $1/2$. We have used the arrow $\longrightarrow$ to indicate one step when changing the matrix. Later on, we will gradually change the matrix and use the arrow $\longrightarrow$, each time an operation is used. Below the arrow, we write which operation is used (in this case $R_1 \leftarrow (1/2) \cdot R_1$).

Finally, the third operation, adding d times equation R_j to an equation R_i (where $i \neq j$ and $d \in \mathbb{F}$), simply means that the linear equation R_i given by $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$ is replaced by the equation $(a_{i1} + da_{j1})x_1 + \cdots + (a_{in} + da_{jn})x_n = b_i + db_j$. One can briefly state this by writing that the linear equation R_i is replaced by $R_i + d \cdot R_j$, or in other words as $R_i \leftarrow R_i + d \cdot R_j$.

|||| Example 6.16

Again let us use the system from Example 6.8 to illustrate the effect of the operation $R_1 \leftarrow R_1 + 2 \cdot R_2$. This amounts to replacing the system

$$\begin{cases} 2 \cdot x_1 & + & 6 \cdot x_2 & + & 0 \cdot x_3 & + & 1 \cdot x_4 & = & 0 \\ 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{cases}$$

by

$$\begin{cases} 10 \cdot x_1 & + & 6 \cdot x_2 & + & (-2) \cdot x_3 & + & 1 \cdot x_4 & = & 12 \\ 4 \cdot x_1 & + & 0 \cdot x_2 & + & (-1) \cdot x_3 & + & 0 \cdot x_4 & = & 6 \end{cases}.$$

In matrix notation, we obtain

$$\begin{bmatrix} 2 & 6 & 0 & 1 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix} \underset{R_1 \leftarrow R_1 + 2 \cdot R_2}{\longrightarrow} \begin{bmatrix} 10 & 6 & -2 & 1 & 12 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix}.$$

Hence the effect of the operation $R_1 \leftarrow R_1 + 2 \cdot R_2$ on the augmented matrix, is that the first row is replaced by the first row plus two times the second row.

As is clear from the examples, the effect of the three operations $R_i \leftrightarrow R_j$, $R_j \leftarrow c \cdot R_j$, and $R_i \leftarrow R_i + d \cdot R_j$ can be seen as easy operations on the rows of the augmented matrix of the system of linear equations we started with. For this reason, they are called *elementary row operations*. This is in fact also the reason why we used capital R in the labels $R_1, \ldots, R_m$ for the linear equations in our system: the R simply was inspired by the first letter in the word "row".

Now let us make sure that when using any of these elementary operations, the solution set of the new system is identical to that of the original system of linear equations. In fact, let us state this as a theorem:

Theorem 6.17

Let $R_1, \ldots, R_m$ be a system of m linear equations in n variables over a field $\mathbb{F}$. Further, let i and j be two distinct integers between 1 and m. Then any of the systems obtained by applying one of the operations $R_i \leftrightarrow R_j$, $R_j \leftarrow c \cdot R_j$, with $c \in \mathbb{F} \setminus \{0\}$ or $R_i \leftarrow R_i + d \cdot R_j$, with $d \in \mathbb{F}$, has the same set of solutions as the original system.

Proof. We only prove the theorem for the elementary operation $R_i \leftarrow R_i + d \cdot R_j$. The reader is encouraged to check that the theorem is also true for the remaining two elementary operations. We need to show that the set of solutions of the system of linear equations $R_1, \ldots, R_{i-1}, R_i, R_{i+1}, \ldots, R_m$ is the same as the set of solutions of the system given by $R_1, \ldots, R_{i-1}, R_i + d \cdot R_j, R_{i+1}, \ldots, R_m$. Let us denote the first set of solutions by S and the second set by T. We wish to show that $S = T$.

First of all, we claim that $S \subseteq T$. Therefore, let us choose arbitrary $(v_1, \ldots, v_n) \in S$. We want to show that $(v_1, \ldots, v_n) \in T$. In other words, assuming that $(v_1, \ldots, v_n) \in \mathbb{F}^n$ is a common solution to the linear equations $R_1, \ldots, R_m$, we need to show that it also is a common solution to the linear equations $R_1, \ldots, R_{i-1}, R_i + d \cdot R_j, R_{i+1}, \ldots, R_m$. But then we only need to show that $(v_1, \ldots, v_n)$ is a solution to $R_i + d \cdot R_j$. This is certainly true, since if $(v_1, \ldots, v_n)$ is a common solution to R_i and R_j, then it is also a solution to $R_i + d \cdot R_j$ for any constant $d \in \mathbb{F}$. Hence $(v_1, \ldots, v_n) \in T$. Since we chose $(v_1, \ldots, v_n) \in S$ arbitrarily, this implies that $S \subseteq T$.

Now we claim that $T \subseteq S$. We choose arbitrary $(v_1, \ldots, v_n) \in T$ and now want to show that $(v_1, \ldots, v_n) \in S$. This means that we may assume that $(v_1, \ldots, v_n) \in \mathbb{F}^n$ is a common solution to the linear equations $R_1, \ldots, R_{i-1}, R_i + d \cdot R_j, R_{i+1}, \ldots, R_m$. We need to show that $(v_1, \ldots, v_n)$ is a solution to R_i. However, this is true, since $R_i = (R_i + d \cdot R_j) - d \cdot R_j$. Hence $(v_1, \ldots, v_n) \in S$. Since we chose $(v_1, \ldots, v_n) \in T$ arbitrarily, this implies that $T \subseteq S$.

Now that we have shown that $S \subseteq T$ and $T \subseteq S$, Lemma 2.6 implies that $S = T$, which is what we wanted to show. □

It turns out that with these three rather elementary operations in hand, we can find the set of solutions to any system of linear equations. Using one elementary row operation, may not simplify a system of linear equation so much, but the idea is that if we use several elementary row operations in succession, we can transform any given system into a much simpler one. In the next sections, we will see how, but for now, let us consider an example.

|||| Example 6.18

Let us revisit Example 6.8. There we considered the following system of 2 equations in 4 variables over $\mathbb{R}$:

$$\begin{cases} 2 \cdot x_1 + 6 \cdot x_2 + 0 \cdot x_3 + 1 \cdot x_4 = 0 \\ 4 \cdot x_1 + 0 \cdot x_2 + (-1) \cdot x_3 + 0 \cdot x_4 = 6 \end{cases}.$$

Let us simplify this system, applying elementary row operations. As we have seen in Theorem 6.17, this does not change the solution set of the system. Since it is much more compact to work with the augmented matrix of the system, let us do that as well.

First, applying the transformation $R_1 \leftarrow (1/2) \cdot R_1$, we obtain the augmented matrix:

$$\begin{bmatrix} 2 & 6 & 0 & 1 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix} \underset{R_1 \leftarrow (1/2) \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix}.$$

The point of this operation, was to get a one in the first entry of the first row. This makes it easy to eliminate the x_1 variable from the second equation. In other words, in the next step we want to create a zero in the first entry of the second row. We achieve this applying the elementary row operation $R_2 \leftarrow R_2 - 4 \cdot R_1$, since then we obtain

$$\begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 4 & 0 & -1 & 0 & 6 \end{bmatrix} \underset{R_2 \leftarrow R_2 - 4 \cdot R_1}{\longrightarrow} \begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & -12 & -1 & -2 & 6 \end{bmatrix}.$$

Now we simplify further, by making the coefficient for x_2 in the second equation equal to one. In other words, we now want to make the second entry in the second row equal to one. To achieve this, we apply $R_2 \leftarrow (-1/12) \cdot R_2$:

$$\begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & -12 & -1 & -2 & 6 \end{bmatrix} \underset{R_2 \leftarrow (-1/12) \cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & 1 & 1/12 & 2/12 & -6/12 \end{bmatrix}.$$

The fractions in the resulting matrix can actually be simplified a bit, so we could also have written:

$$\begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & -12 & -1 & -2 & 6 \end{bmatrix} \underset{R_2 \leftarrow (-1/12) \cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & 1 & 1/12 & 1/6 & -1/2 \end{bmatrix}.$$

The corresponding system is now nearly as simple as we can make it, but we can still use the second equation to get rid of the x_2 term in the first equation using $R_1 \leftarrow R_1 - 3 \cdot R_2$:

$$\begin{bmatrix} 1 & 3 & 0 & 1/2 & 0 \\ 0 & 1 & 1/12 & 1/6 & -1/2 \end{bmatrix} \underset{R_1 \leftarrow R_1 - 3 \cdot R_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & -1/4 & 0 & 3/2 \\ 0 & 1 & 1/12 & 1/6 & -1/2 \end{bmatrix}.$$

The corresponding system of linear equations is:

$$\begin{cases} x_1 + (-1/4) \cdot x_3 = 3/2 \\ x_2 + (1/12) \cdot x_3 + (1/6) \cdot x_4 = -1/2 \end{cases}.$$

It is important to remember that by Theorem 6.17, the set of solutions to this last system, is exactly the same as the set of solutions to the system we started with.

It is easy to find solutions $(v_1, v_2, v_3, v_4) \in \mathbb{R}^4$ to the last system: simply choose $v_3, v_4 \in \mathbb{R}$ as you want, then use the linear equations to solve for v_1 and v_2. For example, if we choose $v_3 = 0$ and $v_4 = 3$, then we find that $v_1 = (1/4)v_3 + 3/2 = 3/2$ and $v_2 = -(1/12)v_3 + (-1/6)v_4 - 1/2 = -1$. Hence $(3/2, -1, 0, 3)$ is a solution to the system. More, and in fact all, solutions can be obtained in this way: choose any value for v_3 and v_4 that you like and determine the corresponding v_1 and v_2 from the equations $v_1 = (1/4)v_3 + 3/2$ and $v_2 = -(1/12)v_3 + (-1/6)v_4 - 1/2$.

This example shows that it can help a great deal to simplify a given system of linear equation first, before trying to solve it.

6.3 The reduced row echelon form of a matrix

We have seen in Example 6.18 that using elementary row operations, can help to describe the solution set of a system of linear equations. What we will do now is to show that this approach always works. Rather than working with systems of linear equations, we will work with the coefficient and augmented matrix of the system. We have seen that if the system consists of m linear equations in n variables, then the coefficient matrix is an $m \times n$ matrix, while the augmented matrix is an $m \times (n+1)$ matrix. The entries in these matrices are from $\mathbb{F}$, the field we are working over. As mentioned before, we will typically work with either $\mathbb{F} = \mathbb{R}$, the real numbers, or $\mathbb{F} = \mathbb{C}$, the complex numbers. The set of all $m \times n$ matrices with entries in $\mathbb{F}$ will be denoted by $\mathbb{F}^{m \times n}$. In formulas, we will typically use bold face letters, such as $\mathbf{A}, \mathbf{B}, \ldots$ for matrices.

We begin by defining a special kind of matrix:

Definition 6.19

Let $\mathbb{F}$ be a field and $\mathbf{A} \in \mathbb{F}^{m \times n}$ an $m \times n$ matrix with entries in $\mathbb{F}$. One says that $\mathbf{A}$ is in *reduced row echelon form*, if all of the following are fulfilled.

1. If a row of the matrix contains only zeros, it appears at the bottom of the matrix. Such rows are called zero rows.
2. The left-most non-zero entry in any non-zero row is equal to 1. This entry is called the *pivot* of the row.
3. Pivots of two non-zero rows of the matrix do not occur in the same column. Moreover, the pivot of the upper row is further to the left than the pivot of the lower row.
4. If a column of the matrix contains a pivot, then all other entries in that column are 0.

A matrix satisfying the first three items, but not necessarily the fourth item, is said to be in *row echelon form*.

Example 6.20

The 1×4 matrices $[0\,0\,0\,0]$ and $[0\,0\,1\,5]$ are both in reduced row echelon form. Also the 2×5 matrix

$$\begin{bmatrix} 1 & 0 & -1/4 & 0 & 3/2 \\ 0 & 1 & 1/12 & 1/6 & -1/2 \end{bmatrix}$$

which we obtained at the end of Example 6.18, is in reduced row echelon form.

An example of a 1×4 matrix that is not in reduced row echelon form is: $[0\,0\,2\,0]$. Indeed, the left-most non-zero entry in the first (and only) row is not equal to 1. An example of a 3×4 matrix that is not in reduced row echelon form is:

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

This matrix is in row echelon form, but not in reduced row echelon form. The problem here is the third column. This column contains a pivot, namely the pivot of the second row, but apart from the pivot, this column contains another non-zero element (the 2).

The reason that reduced row echelon forms are so important for us is the following result:

Theorem 6.21

Let $\mathbf{A} \in \mathbb{F}^{m \times n}$ be a matrix. Then $\mathbf{A}$ can be brought into reduced row echelon form using elementary row operations.

Proof. We will give a sketch of the proof. The strategy is to first bring the matrix in row echelon form, and afterwards in reduced row echelon form. Let us therefore first show that we can use elementary row operations to bring the matrix $\mathbf{A}$ in row echelon form. To do this, we will use induction on m, the number of rows.

If $m = 1$ (the base case of the induction), then the only way $\mathbf{A}$ cannot be in row echelon form, is if the row contains a non-zero entry and the left-most non-zero entry, say c, is not equal to one. Then the operation $R_1 \leftarrow c^{-1} \cdot R_1$ will bring $\mathbf{A}$ in row echelon form.

For the induction step, suppose $m > 1$ is given and that the theorem is true for $(m - 1) \times n$ matrices. If all entries in the matrix $\mathbf{A}$ are zero, it is already in row echelon form (and in fact also reduced row echelon form) and we are done. Therefore, let us now assume that the matrix $\mathbf{A}$ has at least one nonzero entry. We start by choosing the smallest possible j such that the j-th column of $\mathbf{A}$ contains a nonzero entry. In particular, if $j > 1$, then the first $j - 1$ columns of $\mathbf{A}$ are all zero columns. After this, we choose the smallest possible i such that a_{ij}, the (i, j)-th entry of $\mathbf{A}$, is nonzero. Now we perform the operation $R_1 \leftrightarrow R_i$. The first row of the resulting matrix has a nonzero entry in its j-th position, say c, and zero entries in positions 1 up till $j - 1$. Next, we perform the operation $R_1 \leftarrow c^{-1}R_1$, implying that now the j-th entry in the first row has become a 1. If not all elements below this 1 are zeros, we use elementary operations of the form $R_j \leftarrow R_j + dR_1$ for suitably chosen $d \in \mathbb{F}$ to transform the matrix further into a matrix, where there are only zeros below the pivot in row one. We have now transformed the matrix $\mathbf{A}$ into a matrix $\mathbf{B}$ of the form

$$\mathbf{B} = \begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{bmatrix}.$$

In this notation, the first part of the matrix $\mathbf{B}$ was given as

$$\begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}.$$

This reflects the fact that the first $j-1$ columns of $\mathbf{B}$ are zero. The notation is not meant to suggest that the first part of $\mathbf{B}$ contains at least two zero columns. Indeed, if $j=2$, this part just consists of one zero column, since then $j-1=1$. In the case that $j=1$, the first column of the matrix $\mathbf{B}$ is actually not zero at all, but is the column whose first coordinate is 1 and otherwise contains zeroes.

Irrespective on the precise value of j, we now proceed by simply removing the first row of the matrix $\mathbf{B}$ and denote the $(m-1)\times n$ matrix that remains, by $\mathbf{C}$. Using the induction hypothesis, we can conclude that we can use elementary row operations to transform the matrix $\mathbf{C}$ into a matrix $\hat{\mathbf{C}}$ that is in row echelon form. Putting back the first row from $\mathbf{B}$, we find an $m\times n$ matrix, say $\hat{\mathbf{A}}$, that is in row echelon form.

This concludes the inductive proof that any matrix can be brought in row echelon form using elementary row operations. What remains to be done is to bring this matrix in reduced row echelon form. We know by definition of row echelon form that pivots of two non-zero rows of the matrix $\hat{\mathbf{A}}$ do not occur in the same column and moreover, that the pivot of the upper row is further to the left than the pivot of a lower row. Therefore, the entries below a pivot in the matrix $\hat{\mathbf{A}}$, are zero. However, the entries above a pivot in this matrix may not be zero. This can be achieved using elementary row operations of the form $R_i \leftarrow R_i + dR_j$, where row R_j contains a pivot and $i<j$. More precisely, we start using the row containing the right-most pivot to create zeros above this pivot and then work our way to the left, dealing with one pivot at the time. Once we have arrived at the left-most pivot and carried out the sketched procedure for that pivot as well, the obtained matrix will be in reduced row echelon form. □

As an example, we can simply look at Example 6.18. There we used elementary row operations to bring a matrix in reduced row echelon form. There are in principle many different ways to use elementary row operations to transform a given matrix $\mathbf{A}$ into reduced row echelon form. However, for a given matrix $\mathbf{A}$, it turns out that the outcome is always the same. Therefore we can talk about *the* reduced row echelon form of a matrix $\mathbf{A} \in \mathbb{F}^{m\times n}$. In particular, the following definition is justified:

Definition 6.22

Let $\mathbb{F}$ be a field and $\mathbf{A} \in \mathbb{F}^{m\times n}$ a matrix. Then the *rank* of $\mathbf{A}$, denoted by $\rho(\mathbf{A})$, is defined as the number of pivots in the reduced row echelon form of $\mathbf{A}$.

The proof of Theorem 6.21 is very algorithmic in nature and can indeed be made into an algorithm. Let us state the pseudo-code of an algorithm that computes a row echelon form of a matrix. Note how closely it follows the first part of the proof of Theorem 6.21. One could extend the algorithm and obtain pseudo-code of an algorithm that computes the reduced row echelon form of a matrix, but we will not do that.

Algorithm 1 for computing a row echelon form of a matrix

Input: Positive integers m, n and an $m \times n$ matrix $\mathbf{A} \in \mathbb{F}^{m\times n}$
Output: $\mathrm{ref}(\mathbf{A})$, the reduced row echelon form of $\mathbf{A}$

if $\mathbf{A} = \mathbf{0}$ **then**
    $\mathrm{ref}(\mathbf{A}) \leftarrow \mathbf{0}$,
if $m = 1$ and $\mathbf{A} \neq \mathbf{0}$ **then**
    $j \leftarrow$ smallest column index such that $\mathbf{A}_{1j} \neq 0$
    $\mathrm{ref}(\mathbf{A}) \leftarrow (\mathbf{A}_{1j})^{-1} \cdot \mathbf{A}$
if $m > 1$ and $\mathbf{A} \neq \mathbf{0}$ **then**
    $j \leftarrow$ least ℓ such that some row of $\mathbf{A}$ has non-zero ℓ-th entry
    $i \leftarrow$ least i such that the ith row of $\mathbf{A}$ has nonzero j-th entry
    $\mathbf{B} \leftarrow$ the matrix obtained from $\mathbf{A}$ by applying $R_1 \leftrightarrow R_i$
    $b \leftarrow$ the ith entry of the first row of $\mathbf{B}$
    $\mathbf{B} \leftarrow$ the matrix obtained from $\mathbf{B}$ by applying $R_1 \leftarrow b^{-1} \cdot R_1$
    $\mathbf{r} \leftarrow$ the 1-st row of $\mathbf{B}$
    for $i = 2...m$ **do**
        $b \leftarrow$ the first entry of the i-th row of $\mathbf{B}$
        $\mathbf{B} \leftarrow$ the matrix obtained from $\mathbf{B}$ by applying $R_i \leftarrow R_i - bR_1$
    $\mathbf{C} \leftarrow$ the matrix obtained from $\mathbf{B}$ by deleting the first row
    $\mathbf{C} \leftarrow \mathrm{ref}(\mathbf{C})$ (here the algorithm call itself recursively)
    $\mathrm{ref}(\mathbf{A}) \leftarrow$ the matrix obtained by adding $\mathbf{r}$ on top of $\mathbf{C}$

6.4 Computing all solutions to systems of linear equations

Up till now, we have usually written elements from $\mathbb{F}^n$ as n-tuples $(a_1, \ldots, a_n)$. It is quite common to identify $\mathbb{F}^n$ with $\mathbb{F}^{n\times 1}$, that is to say, to identify an n-tuple with an

$n \times 1$ matrix. Such a matrix only contains one column. This means for example that:

$$(1,2,4,7) \text{ is identified with } \begin{bmatrix} 1 \\ 2 \\ 4 \\ 7 \end{bmatrix}.$$

A small warning is in place. Even though, we will always identify $\mathbb{F}^n$ and $\mathbb{F}^{n\times 1}$, some books prefer to identify $\mathbb{F}^n$ and $\mathbb{F}^{1\times n}$.

When performing elementary row operations, we have at times multiplied rows of a matrix with an element c from $\mathbb{F}$ or added one row to another. A similar operation can be performed on columns in a matrix. In particular, it is customary to define

$$c \cdot \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} c \cdot a_1 \\ \vdots \\ c \cdot a_n \end{bmatrix} \text{ and } \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} a_1' \\ \vdots \\ a_n' \end{bmatrix} = \begin{bmatrix} a_1 + a_1' \\ \vdots \\ a_n + a_n' \end{bmatrix}.$$

This notation, combined with the theory of reduced row echelon matrices, will make it possible to determine whether or not a given system of linear equations has solutions, and if yes, to write all solutions down in a systematic way. Let us start with determining when a system has a solution.

Theorem 6.23

Let a system of m linear equations in n variables over a field $\mathbb{F}$ be given. Denote by $\mathbf{A}$ the coefficient matrix of the system and by $[\mathbf{A}|\mathbf{b}]$ its augmented matrix. Then the system has no solution if $\mathbf{A}$ and $[\mathbf{A}|\mathbf{b}]$ do not have the same rank.

Proof. We know from Theorem 6.21, that there exists a sequence of elementary row operations that brings the matrix $\mathbf{A}$ in its row reduced echelon form, say $\hat{\mathbf{A}}$. Since the first n columns of the augmented matrix $[\mathbf{A}|\mathbf{b}]$ are identical with those of the coefficient matrix $\mathbf{A}$, applying exactly the same elementary row operations on $[\mathbf{A}|\mathbf{b}]$ yields a matrix, say $\mathbf{B}$, whose first n columns are identical with those of the reduced row echelon form of $\mathbf{A}$. Therefore we can write $\mathbf{B} = [\hat{\mathbf{A}}|\hat{\mathbf{b}}]$ for some $\hat{\mathbf{b}} \in \mathbb{F}^m$. Let us denote the bottom entry $\hat{\mathbf{b}}$ by $\hat{b}_m$. If the bottom row of $\hat{\mathbf{A}}$ contains a pivot, then the matrix $[\hat{\mathbf{A}}|\hat{\mathbf{b}}]$ is in reduced row echelon form. But then we see that the matrices $\mathbf{A}$ and $[\mathbf{A}|\mathbf{b}]$ have the same rank, contrary to the assumption given in the theorem that $\mathbf{A}$ and $[\mathbf{A}|\mathbf{b}]$ do not have the same rank. Therefore we may assume that the bottom row of $\hat{\mathbf{A}}$ does not contain a pivot, which simply means that this row is the zero row. If the last row of $\hat{\mathbf{A}}$ does not contain a pivot and $\hat{b}_m = 0$, then the matrix $[\hat{\mathbf{A}}|\hat{\mathbf{b}}]$ is in reduced row echelon form and we can conclude that $\rho(\mathbf{A}) = \rho([\mathbf{A}|\mathbf{b}])$, again leading to a contradiction. Therefore we may assume that the bottom row of $\hat{\mathbf{A}}$ does not contain a pivot and that $\hat{b}_m \neq 0$. But then the bottom row of the matrix $[\hat{\mathbf{A}}|\hat{\mathbf{b}}]$ corresponds to the equation $0 \cdot x_1 + \cdots 0 \cdot x_m = \hat{b}_m$. Since this equation has no solution, Theorem 6.17 implies that the system we started with has no solution either. $\square$

|||| Example 6.24

As in Example 6.11, consider the following system of two linear equations in two variables over $\mathbb{R}$:

$$\begin{cases} x_1 & + & x_2 & = & 1 \\ x_1 & + & x_2 & = & 0 \end{cases}.$$

We have already seen in Example 6.11 that this system has no solutions. Let us now try to confirm this using Theorem 6.23. The augmented matrix $[\mathbf{A}|\mathbf{b}]$ is given by

$$[\mathbf{A}|\mathbf{b}] = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Applying the row operation $R_1 \leftrightarrow R_2$ followed by $R_2 \leftarrow R_2 - R_1$, we find the reduced row echelon form of the augmented matrix:

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

Hence $\rho([\mathbf{A}|\mathbf{b}]) = 2$. The reduced row echelon form of the coefficient matrix is the matrix

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

which can be obtained from $\mathbf{A}$ by applying the operation $R_2 \leftarrow R_2 - R_1$. Hence $\rho(\mathbf{A}) = 1$. Since $\rho(\mathbf{A}) \neq \rho([\mathbf{A}|\mathbf{b}])$, Theorem 6.23 implies that indeed the system we started with does not have a solution.

In case $\mathbf{A}$ and $[\mathbf{A}|\mathbf{b}]$ do have the same rank, we can use the theory of reduced row echelon matrices, to describe a solution explicitly. Let us look at a concrete example first.

|||| Example 6.25

Let us consider a system of three linear equations in four variables over $\mathbb{R}$, whose augmented matrix already is in reduced row echelon form:

$$\left\{\begin{array}{ccccccccc} x_1 & + & 2\cdot x_2 & + & & & 3\cdot x_4 & = & 5 \\ & & & & x_3 & + & 4\cdot x_4 & = & 6 \\ & & & & & & 0 & = & 0 \end{array}\right. .$$

We can see that in this case the coefficient matrix $\mathbf{A}$ and augmented matrix $[\mathbf{A}|\mathbf{b}]$ are

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \text{ respectively } [\mathbf{A}|\mathbf{b}] = \begin{bmatrix} 1 & 2 & 0 & 3 & 5 \\ 0 & 0 & 1 & 4 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since both are already in reduced row echelon form, we can immediately determine the ranks of these matrices and conclude that $\rho(\mathbf{A}) = \rho([\mathbf{A}|\mathbf{b}]) = 2$. Theorem 6.23 does therefore not apply, and we cannot conclude anything about the existence of solutions yet. However, a solution is easily determined in the following way: first rewrite the equations in the following way:

$$\left\{\begin{array}{ccc} x_1 & = & 5 - 2\cdot x_2 - 3\cdot x_4 \\ x_3 & = & 6 - 4\cdot x_4 \end{array}\right. .$$

Now we can choose $x_2 = v_2$ and $x_4 = v_4$ as we want for any $v_2, v_4 \in \mathbb{R}$ and then compute the resulting values for x_1 and x_3. For example, choosing $v_2 = v_4 = 0$, we find the solution

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 6 \\ 0 \end{bmatrix}.$$

Exactly the same approach can be used in general to find a solution to a system of linear equations, provided that coefficient and augmented matrix have the same rank. The result is the following:

Theorem 6.26

Let a system of m linear equations in n variables over a field $\mathbb{F}$ be given. Denote by $\mathbf{A}$ the coefficient matrix of the system and by $[\mathbf{A}|\mathbf{b}]$ its augmented matrix and suppose that these matrices have the same rank ρ. Moreover, assume that the pivots of the reduced row echelon form of $\mathbf{A}$ are at the positions $(1, j_1), \ldots, (\rho, j_\rho)$, and that the top ρ entries of the last column of the reduced row echelon form of $[\mathbf{A}|\mathbf{b}]$ are given by $\hat{b}_1, \ldots, \hat{b}_\rho$. Then the m-tuple $(v_1, \ldots, v_n)$ defined as

$$v_j = \begin{cases} \hat{b}_\ell & \text{if } j = j_\ell \text{ for some } \ell = 1, \ldots, \rho, \\ 0 & \text{otherwise.} \end{cases}$$

is a possible solution to the system.

Proof. The idea of the proof is simply to generalize the approach used in Example 6.25. First of all, we use the equations corresponding to the rows of the reduced row echelon form of the augmented matrix $[\mathbf{A}|\mathbf{b}]$ to express the variables x_j with $j \in \{j_1, \ldots, j_\ell\}$ in terms of the remaining $n - \rho$ variables. Then putting all these remaining variables $x_j, j \notin \{j_1, \ldots, j_\ell\}$ equal to zero, we find that $x_j = \hat{b}_\ell$ for $j = j_\ell$ and $\ell = 1, \ldots, \rho$. Hence the n-tuple $(v_1, \ldots, v_n)$ is indeed a solution to the system whose augmented matrix is the reduced row echelon form of $[\mathbf{A}|\mathbf{b}]$. Now applying Theorem 6.17, we see that this n-tuple is also a solution to the system we started with. $\square$

Theorem 6.26 does by no means state that the indicated solution is the only solution. Indeed, we know from Theorem 6.10 that there can be more. Recall that a solution to an inhomogeneous system of linear equations was called a particular solution. If the system of linear equations in inhomogeneous, Theorem 6.26 therefore gives such a particular solution, provided it exists.

Corollary 6.27

Let a system of m linear equations in n variables over a field $\mathbb{F}$ be given. Denote by $\mathbf{A}$ the coefficient matrix of the system and by $[\mathbf{A}|\mathbf{b}]$ its augmented matrix. Then the system has no solution if and only if $\mathbf{A}$ and $[\mathbf{A}|\mathbf{b}]$ do not have the same rank.

Proof. The "if" part is precisely Theorem 6.23. In other words, we have already seen in

Theorem 6.23 that if $\rho(\mathbf{A}) \neq \rho([\mathbf{A}|\mathbf{b}])$, then the system has no solutions. Conversely, if $\rho(\mathbf{A}) = \rho([\mathbf{A}|\mathbf{b}])$, then Theorem 6.26 implies that the system does have at least one solution. □

With Corollary 6.27 we can determine exactly if a given system of linear equations has a solution. Moreover, using Theorem 6.26, we can determine at least one solution if such solutions exist. Now recall that in Theorem 6.10, we have seen that in order to find all solutions of an inhomogeneous system of linear equations, it is enough to find all solutions of the corresponding homogeneous system of linear equations and one particular solution of the inhomogeneous system. Therefore, what is left to do, is to describe how one finds all solutions to a homogeneous system of linear equations. This is precisely the aim of the next theorem, but let us look at an example first to get the idea.

|||| Example 6.28

Let us consider a system of three linear equations in four variables over $\mathbb{R}$, whose augmented matrix already is in reduced row echelon form:

$$\left\{ \begin{array}{ccccccccc} x_1 & + & 2 \cdot x_2 & + & & & 3 \cdot x_4 & = & 0 \\ & & & & x_3 & + & 4 \cdot x_4 & = & 0 \\ & & & & & & 0 & = & 0 \end{array} \right. .$$

This system is similar to the system of linear equation we studied in Example 6.25, but this time it is homogeneous. In particular, the coefficient matrices of the system above and the system from Example 6.25 are the same and as observed in Example 6.25, it is in reduced row echelon form.

It is not hard to find all solutions to the system. Since the coefficient matrix of the system is in reduced row echelon form with pivots in the first and third column, we can express x_1 and x_3 in terms of x_2 and x_4. More concretely, we can rewrite the equations as

$$\left\{ \begin{array}{ccc} x_1 & = & -2 \cdot x_2 - 3 \cdot x_4 \\ x_3 & = & -4 \cdot x_4 \end{array} \right. .$$

Hence any solution $(v_1, v_2, v_3, v_4) \in \mathbb{R}^4$ to the system satisfies

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} -2 \cdot v_2 - 3 \cdot v_4 \\ v_2 \\ -4 \cdot v_4 \\ v_4 \end{bmatrix} = v_2 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + v_4 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix}.$$

Therefore, we can think of $v_2, v_4 \in \mathbb{R}$ as parameters that we can choose arbitrarily, each choice giving us a solution to the system of linear equations we started with. Changing notation from v_2 to t_1 and v_4 to t_2, we see that any solution to the system is of the form

$$t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \qquad (t_1, t_2 \in \mathbb{R})$$

Conversely, since a direct check shows that

$$\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$

are solution to the system, Theorem 6.9 implies that for any $t_1, t_2 \in \mathbb{R}$, the expression

$$t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$

is also a solution. Putting this together, we see that the solutions to the homogeneous system of linear equations we started are precisely those $(v_1, v_2, v_3, v_4) \in \mathbb{R}^4$ such that

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \qquad (t_1, t_2 \in \mathbb{R}).$$

One calls such a description of the solutions, the *general solution* of the homogeneous system. The solution set to the homogeneous system of linear equations

$$\left\{ \begin{array}{ccccccccc} x_1 & + & 2 \cdot x_2 & + & & & 3 \cdot x_4 & = & 0 \\ & & & & x_3 & + & 4 \cdot x_4 & = & 0 \\ & & & & & & 0 & = & 0 \end{array} \right.$$

is precisely given by

$$\left\{ t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \;\middle|\; t_1, t_2 \in \mathbb{R} \right\}.$$

In this example, we started out with a homogeneous system of linear equations whose coefficient matrix was in reduced row echelon form. This was the reason that we could determine all solutions relatively fast. From the previous sections, we know however that even if we start with a more complicated system, we can always use elementary row operations to transform it in such a way that the resulting coefficient matrix is in reduced echelon form. Basically, Example 6.28 describes how to find all solutions, once the coefficient matrix of the system of linear equations is in reduced row echelon form. Exactly the same ideas work for any homogeneous system of linear equations: first simplify the system by bringing its coefficient matrix in reduced row echelon form, then follow the procedure exemplified in Example 6.28. It is possible to describe the outcome for the general case and for the sake of completeness we do so in the following theorem. However, when asked to solve a homogeneous system of linear equations in practice, it is often easier not to use this theorem, but instead to use a procedure similar to the one in Example 6.28 directly.

|||| Theorem 6.29

Let a homogeneous system of m linear equation in n variables over a field $\mathbb{F}$ be given. Denote the coefficient matrix of this system by $\mathbf{A}$ and let $\hat{\mathbf{A}}$ denote the reduced row echelon form of $\mathbf{A}$. Further, suppose that $\hat{\mathbf{A}}$ has ρ pivots in columns $j_1, \ldots, j_\rho$, and denote by

$$\mathbf{c}_1 = \begin{bmatrix} c_{11} \\ \vdots \\ c_{m1} \end{bmatrix}, \ldots, \mathbf{c}_{n-\rho} = \begin{bmatrix} c_{1\,n-\rho} \\ \vdots \\ c_{m\,n-\rho} \end{bmatrix}$$

the $n-\rho$ columns of $\hat{\mathbf{A}}$ not containing a pivot. Finally, define

$$\mathbf{v}_1 = \begin{bmatrix} v_{11} \\ \vdots \\ v_{n1} \end{bmatrix}, \ldots, \mathbf{v}_{n-\rho} = \begin{bmatrix} v_{1\,n-\rho} \\ \vdots \\ v_{n\,n-\rho} \end{bmatrix}$$

by

$$v_{j\,i} = \begin{cases} -c_{\ell i} & \text{if } j = j_\ell \text{ for some } \ell = 1, \ldots, \rho, \\ 1 & \text{if } \mathbf{c}_i \text{ is the } j\text{-th column in } \hat{\mathbf{A}}, \\ 0 & \text{otherwise.} \end{cases}$$

Then the solution set of the given homogeneous system of linear equations is given by

$$\left\{ t_1 \cdot \begin{bmatrix} v_{11} \\ \vdots \\ v_{n1} \end{bmatrix} + \cdots + t_{n-\rho} \cdot \begin{bmatrix} v_{1\,n-\rho} \\ \vdots \\ v_{n\,n-\rho} \end{bmatrix} \mid t_1, \ldots, t_{n-\rho} \in \mathbb{F} \right\}.$$

Proof. We will not prove this theorem, but only indicate the idea of the proof. First of all Theorem 6.17 is used to conclude that the homogeneous system with coefficient matrix $\mathbf{A}$ has exactly the same solution set as the homogeneous system with coefficient matrix $\hat{\mathbf{A}}$. Then the same approach as in Example 6.28 is used to describe all solutions to the homogeneous system with coefficient matrix $\hat{\mathbf{A}}$. $\square$

The expression

$$t_1 \cdot \begin{bmatrix} v_{11} \\ \vdots \\ v_{n1} \end{bmatrix} + \cdots + t_{n-\rho} \cdot \begin{bmatrix} v_{1\,n-\rho} \\ \vdots \\ v_{n\,n-\rho} \end{bmatrix} \quad (t_1, \ldots, t_{n-\rho} \in \mathbb{F})$$

is called the *general solution* of the homogeneous system with coefficient matrix $\mathbf{A}$. Looking back at Example 6.28, we see that the general solution of the homogeneous system of linear equations studied in that example was shown to be equal to

$$t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \quad (t_1, t_2 \in \mathbb{R}).$$

|||| Corollary 6.30

Let a homogeneous system of m linear equation in n variables over a field $\mathbb{F}$ be given. Denote the coefficient matrix of this system by $\mathbf{A}$. Then the homogeneous system has only the all-zero tuple $(0, \ldots, 0) \in \mathbb{F}^n$ as solution if and only if $\rho(\mathbf{A}) = n$.

Proof. Theorem 6.29 implies that if the rank of $\mathbf{A}$ is less than n, then there exists a nonzero solution. Conversely, if the rank of $\mathbf{A}$ is equal to n, the number of parameters t_i in the description of the solution set in Theorem 6.29, is zero. This means that only the all-zero tuple $(0, \ldots, 0)$ is a solution. $\square$

The status is now that we can determine all solutions to any homogeneous system of linear equations (which we called the general solution of the homogeneous system), can determine whether or not an inhomogeneous system has a solution, and find such a solution (which we called a particular solution) if it does. Hence using Theorem 6.10, we can in this case also determine a formula describing all solutions to an inhomogeneous system of linear equations: it is simply the sum of a particular solution and the general solution of the corresponding homogeneous system. This sum is called the *general solution* of the inhomogeneous system. Therefore we have answered in a constructive way all three questions posed at the end of Section 6.1.

Let us finish this section with an example, where we compute the general solution of an inhomogeneous system of linear equations.

|||| Example 6.31

Let us return to the inhomogeneous system of linear equations considered in Example 6.25:

$$\begin{cases} x_1 + 2\cdot x_2 + 3\cdot x_4 = 5 \\ x_3 + 4\cdot x_4 = 6 \\ 0 = 0 \end{cases}.$$

We have computed a particular solution in Example 6.25 and the general solution of the corresponding homogeneous system in Example 6.28. Using these previous calculations in combination with Theorem 6.10, we conclude that the general solution of the inhomogeneous system is given by:

$$\begin{bmatrix} 5 \\ 0 \\ 6 \\ 0 \end{bmatrix} + t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \quad (t_1, t_2 \in \mathbb{R}).$$

The solution set of the inhomogeneous system is therefore:

$$\left\{ \begin{bmatrix} 5 \\ 0 \\ 6 \\ 0 \end{bmatrix} + t_1 \cdot \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \cdot \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \mid t_1, t_2 \in \mathbb{R} \right\}.$$

6.5 Uniqueness of the reduced row echelon form

Previously, we have stated that a given matrix $\mathbf{A} \in \mathbb{F}^{m\times n}$ has a unique reduced row echelon form. Existence was shown in Theorem 6.21 and in this section we want to show uniqueness. This section can be skipped and is only meant for the reader who wants to see a proof of the uniqueness of the reduced row echelon form.

|||| Theorem 6.32

Let $\mathbb{F}$ be a field and $\mathbf{A} \in \mathbb{F}^{m \times n}$ a matrix. Suppose that $\mathbf{A}$ can be transformed using a sequence of elementary row operations to a matrix $\mathbf{B}_1$ in reduced row echelon form, but using another sequence of elementary row operations to a matrix $\mathbf{B}_2$ in reduced row echelon form. Then $\mathbf{B}_1 = \mathbf{B}_2$.

Proof. From Theorem 6.17, we know that the homogeneous systems of linear equations with coefficient matrices $\mathbf{A}$, $\mathbf{B}_1$, and $\mathbf{B}_2$ all have exactly the same solutions. The idea of the proof is to show that the homogeneous systems of linear equations with coefficient matrices $\mathbf{B}_1$ and $\mathbf{B}_2$ only can have the same solutions if $\mathbf{B}_1 = \mathbf{B}_2$. Moreover, we use induction on n, the number of columns.

Let us start with the induction basis. If $n = 1$, there are only two possible reduced row echelon forms: the $m \times 1$ matrices

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The first can only be a reduced row echelon form of $\mathbf{A}$, if $\mathbf{A}$ was the zero $m \times 1$ matrix to begin with. Performing any elementary row operation on the zero matrix, results in the zero matrix again. Hence if $\mathbf{B}_1$ or $\mathbf{B}_2$ is the zero matrix, then $\mathbf{A} = \mathbf{B}_1 = \mathbf{B}_2$, since they are all equal to the zero matrix. Now suppose that $\mathbf{B}_1$ or $\mathbf{B}_2$ is equal to the second possible $m \times 1$ reduced row echelon matrix. If $\mathbf{B}_1 \neq \mathbf{B}_2$, then at least one of them is equal to the only other $m \times 1$ reduced row echelon form matrix, namely the zero matrix. But we have just seen that this would imply that both $\mathbf{B}_1$ and $\mathbf{B}_2$ are equal to the zero matrix. This contradiction shows that if $\mathbf{B}_1$ or $\mathbf{B}_2$ is equal to the second $m \times 1$ reduced row echelon matrix, then $\mathbf{B}_1 = \mathbf{B}_2$.

We continue to the induction step. Assume $n > 1$ and that the theorem is true for $n - 1$. For any $m \times n$ matrix $\mathbf{A}$, let us denote by $\mathbf{A}|_{n-1}$, the $m \times (n-1)$ matrix one obtains by removing the final column of $\mathbf{A}$. The induction hypothesis implies that $\mathbf{A}|_{n-1}$ has a unique reduced row echelon form. Moreover, if $\mathbf{B}$ is an $m \times n$ matrix in reduced row echelon form, then also the matrix $\mathbf{B}|_{n-1}$ is in reduced row echelon form. This implies that if $\mathbf{B}_1$ and $\mathbf{B}_2$ are two possible reduced row echelon forms of $\mathbf{A}$, then the induction hypothesis implies that $\mathbf{B}_1|_{n-1} = \mathbf{B}_2|_{n-1}$. In other words: the first $n-1$ columns of $\mathbf{B}_1$ $\mathbf{B}_2$ are identical. Only the n-th (i.e., the last) columns may be distinct. Now denote by ρ the number of pivots occurring in $\mathbf{B}_1|_{n-1}$. If the n-th column of $\mathbf{B}_1$ contains a pivot, this column contains zeros only, except in the $(\rho+1)$-th position, where it contains a one. Hence any solution $(v_1,\ldots,v_n) \in \mathbb{F}^n$ to the homogeneous system of linear equations with coefficient matrix $\mathbf{B}_1$, satisfies $v_n = 0$. Conversely, using Theorem 6.29, if the n-th column of $\mathbf{B}_1$ does not contain a pivot, there exists a solution $(v_1,\ldots,v_n)$ such that $v_n = 1$. A similar reasoning applies to the last column of $\mathbf{B}_2$. Using Theorem 6.17, we can however conclude that the homogeneous systems of linear equations with coefficient matrices $\mathbf{B}_1$, $\mathbf{B}_2$, and $\mathbf{A}$ all have exactly the same solution sets. It follows that either a pivot occurs in the n-th columns of both $\mathbf{B}_1$ and $\mathbf{B}_2$, or that no pivot occurs in the n-th columns of both $\mathbf{B}_1$ and $\mathbf{B}_2$. In the first case, we have already seen that the n-th columns are completely determined, implying that $\mathbf{B}_1 = \mathbf{B}_2$. In the second case, we can conclude that there is exactly one solution to the homogeneous system of linear equations with coefficient matrix $\mathbf{A}$ that has a zero in all variables corresponding to the columns not containing pivots, except in the n-th column, where it has a one. Using Theorem 6.29, one sees that the coefficients of this solution completely determine the n-th column of a reduced row echelon form of $\mathbf{A}$. We conclude that $\mathbf{B}_1 = \mathbf{B}_2$ also in the second case where no pivot occurs in the n-th columns of both $\mathbf{B}_1$ and $\mathbf{B}_2$. $\square$