LINEAR ALGEBRA

GABRIEL NAGY

Contents

1. Systems of linear equations

Linear Algebra is the branch of mathematics concerned with the study of systems of linear equations, vectors and vector spaces, and linear transformations. The equations are called a system when there is more than one equation, and they are called linear when the unknown appears as a multiplicative factor with power zero or one. Systems of linear equations is the main subject of this Section, and an example is given by Eqs. (1.3)-(1.4). An example of a vector is an oriented segment, which may belong to the real line $\mathbb{R}$, or to the plane $\mathbb{R}^2$, or to space $\mathbb{R}^3$. These three sets, together with a preferred point that is called the origin, are examples of vector spaces (see Fig. 1). Elements in these spaces are oriented segments with origin at the origin point and head at any point in these spaces. The origin of the word "space" in the term "vector space" originates precisely in these first examples, which were associated with the physical space. Two operations are defined on oriented segments: An oriented segment can be stretched or compressed, and two oriented segments with the same origin point can be added using the parallelogram law. An addition of several stretched or compressed vectors is called a linear combination. Linear transformations are a particular type of functions on vectors that preserve the operation of linear combination. This is the essential structure called vector space. These notes are meant to be an elementary introduction into this subject.

1.1. **Row picture.** A central problem in linear algebra is to find solutions of a system of linear equations. A 2×2 **linear system** is a system of two linear

Date: September 27, 2007.

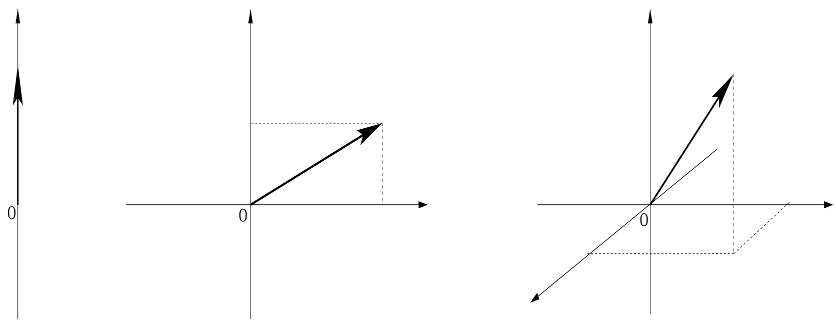

FIGURE 1. Example of vectors in the line, plane and space, respectively.

equations in two unknowns, that is, given the real numbers a_{11}, a_{12}, a_{21}, a_{22}, b_1, and b_2, find the real numbers x and y solutions of

$$a_{11}x + a_{12}y = b_1, \tag{1.1}$$

$$a_{21}x + a_{22}y = b_2. \tag{1.2}$$

These equations are called a system because there is more than one equation, and they are called linear because the unknown appears as a multiplicative factor with power zero or one. An example of a linear system is the following: Find the numbers x and y solutions of

$$2x - y = 0, \tag{1.3}$$

$$-x + 2y = 3. \tag{1.4}$$

The **row picture** consists of finding the solutions to the system as the intersection of all solutions to every single equation of the system. The individual equations are called **row equations**, or simply row of the system. The solution is found geometrically in Fig. 2. Analytically, the solution can be found by substitution:

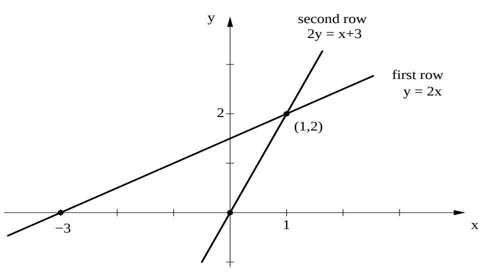

FIGURE 2. The solution of a 2×2 linear system in the row picture is the intersection of the two lines, which are the solutions of each row equation.

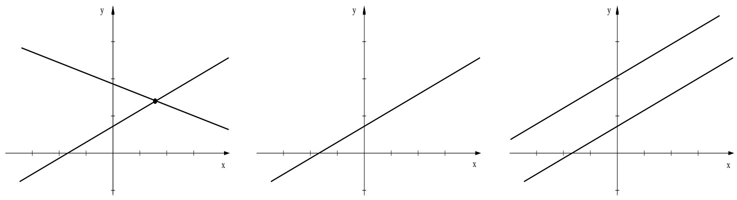

FIGURE 3. An example of the cases given in *(i)-(iii)*, in Theorem 1.

$$2x - y = 0 \quad \Rightarrow \quad y = 2x \quad \Rightarrow \quad -x + 4x = 3 \quad \Rightarrow \quad \begin{cases} x = 1, \\ y = 2. \end{cases}$$

A consequence of the row picture in 2×2 linear systems is the following result.

Theorem 1. *Every 2×2 linear system satisfies only one of the following statements:*

(i) There exists a unique solution;
(ii) There exist infinity many solutions.
(iii) There exists no solution;

Proof of Theorem 1: The solutions of each equation in a 2×2 linear system represents a line in $\mathbb{R}^2$. Two lines in $\mathbb{R}^2$ can intersect at a point, or can be coincident, or can be parallel but not coincident. These are the cases given in *(i)-(iii)*, respectively, and they are represented geometrically in Fig. 3. □

An $m \times n$ **linear system** is defined as a system of m linear equations in n unknowns, that is, given the real numbers a_{ij} and b_i, with $i = 1, \cdots, m$ and $j = 1, \cdots n$, find the real numbers x_j solutions of

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

An $m \times n$ linear system is called **consistent** iff it has solutions, and it is called **inconsistent** iff it has no solutions. Examples of a 2×3 and 3×3 linear systems are given, respectively, as follows,

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ -3x_1 + x_2 - \frac{1}{3}x_3 &= 6 \end{aligned} \qquad \begin{aligned} 2x_1 + x_2 + x_3 &= 2 \\ -x_1 + 2x_2 &= 1 \\ x_1 - x_2 + 2x_3 &= -2. \end{aligned} \tag{1.5}$$

The row picture is appropriate to solve small systems of linear equations. However it becomes difficult to carry out in 3×3 and bigger systems. For example, find the numbers x_1, x_2, x_3 solutions of the 3×3 linear system above. Substitute the second equation into the first,

$$x_1 = -1 + 2x_2 \quad \Rightarrow \quad x_3 = 2 - 2x_1 - x_2 = 2 + 2 - 4x_2 - x_2 \quad \Rightarrow \quad x_3 = 4 - 5x_2;$$

then, substitute the second equation and $x_3 = 4 - 5x_2$ into the third equation,

$$(-1 + 2x_2) - x_2 + 2(4 - 5x_2) = -2 \quad \Rightarrow \quad x_2 = 1,$$

and then, substituting backwards, $x_1 = 1$ and $x_3 = -1$, so the solution is a single point in space $(1, 1, -1)$.

Graphically, the solution of each separate equation represents a plane in $\mathbb{R}^3$. A solution to the system is a point that below to the three planes. In the example above there is a unique solution, the point $(1, 1, -1)$, which means that the three planes intersect at a single point. In the general case, a 3×3 system can have a unique solution, infinitely many solutions or no solutions at all, depending on how the three planes in space intersect among them. The case with unique solution was represented in Fig. 4, while two possible situations corresponding to no solution are given in Fig. 5. Finally, two cases of 3×3 linear system having infinitely many solutions are pictured in Fig 6, where in the first case the solutions form a line, and in the second case the solution form a plane because the three planes coincide.

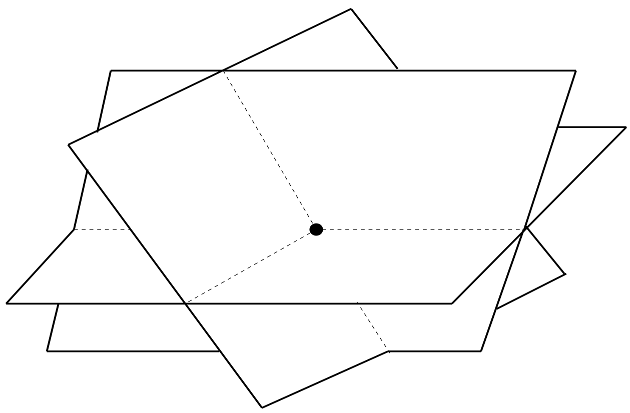

FIGURE 4. Planes representing the solutions of each row equation in a 3×3 linear system having a unique solution.

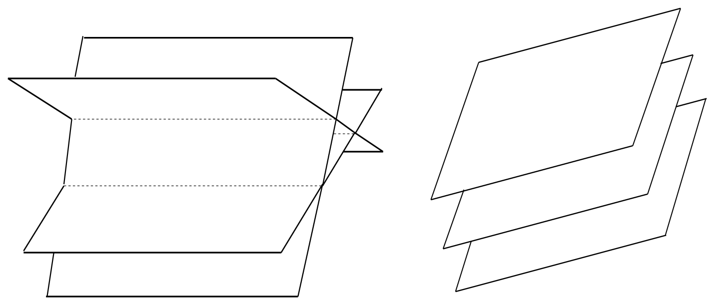

FIGURE 5. Two cases of planes representing the solutions of each row equation in 3×3 linear systems having no solutions.

The solutions to bigger than a 3×3 linear system can not be represented graphically, and the substitution method becomes more involved to solve, hence alternative ideas are needed to solve such systems. In the next section we introduce

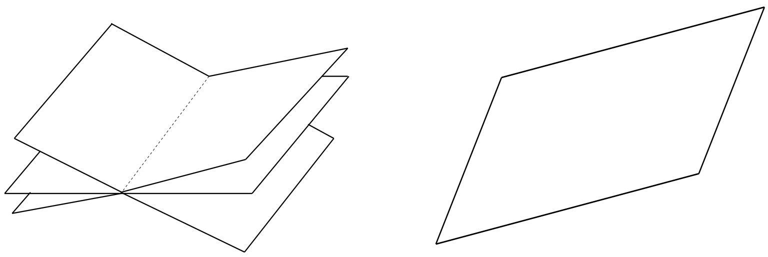

FIGURE 6. Two cases of planes representing the solutions of each row equation in 3×3 linear systems having infinity many solutions.

the column picture, which together with Gauss elimination operations prove to be appropriate to solve efficiently large systems of linear equations.

1.2. **Column picture.** Consider again the linear system in Eqs. (1.3)-(1.4) and introduce a change in the names of the unknowns, calling them x_1 and x_2 instead of x and y. The problem is to find the numbers x_1, and x_2 solutions of

$$2x_1 - x_2 = 0, \tag{1.6}$$

$$-x_1 + 2x_2 = 3. \tag{1.7}$$

We know that the answer is $x_1 = 1$, $x_2 = 2$. The row picture consisted in solving each row separately. The main idea in the column picture is to interpret the 2×2 linear system as an addition of new objects, in the following way,

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix} x_1 + \begin{bmatrix} -1 \\ 2 \end{bmatrix} x_2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}. \tag{1.8}$$

We call these new objects column vectors, and we use boldface letters to denote them, that is,

$$\boldsymbol{a}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \boldsymbol{a}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

We can represent these vectors in the plane, as it is shown in Fig. 7. This column

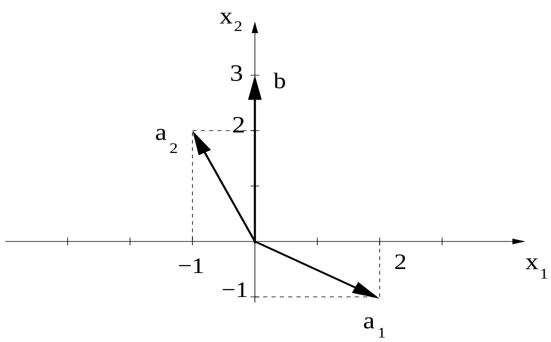

FIGURE 7. Graphical representation of column vectors in the plane.

vector interpretation of a 2×2 linear system determines the addition law of vectors and the multiplication law of a vector by a number. In the example above, we know that the solution is given by $x_1 = 1$ and $x_2 = 2$, therefore in the column picture interpretation the following equation must hold

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix} 2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

This example and the study of other examples determines the multiplication law of a vector by numbers and the addition law of two vectors, according the following equations,

$$\begin{bmatrix} -1 \\ 2 \end{bmatrix} 2 = \begin{bmatrix} (-1)2 \\ (2)2 \end{bmatrix}, \qquad \begin{bmatrix} 2 \\ -1 \end{bmatrix} + \begin{bmatrix} -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 2-2 \\ -1+4 \end{bmatrix}.$$

The study of several examples of 2×2 linear systems in the column picture determines the following rule. Given any 2-vectors $\boldsymbol{u} = \left[\begin{smallmatrix} u_1 \\ u_2 \end{smallmatrix}\right]$ and $\boldsymbol{v} = \left[\begin{smallmatrix} v_1 \\ v_2 \end{smallmatrix}\right]$, and real numbers a and b, introduce the **linear combination** of $\boldsymbol{u}$ and $\boldsymbol{v}$ as follows,

$$a \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + b \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} au_1 + bv_1 \\ au_2 + bv_2 \end{bmatrix}.$$

A linear combination includes the particular cases of addition ($a = b = 1$), and multiplication of a vector by a number ($b = 0$), respectively given by,

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}, \qquad a \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} au_1 \\ au_2 \end{bmatrix}.$$

The addition law in terms of components is represented graphically by the parallelogram law, as it can be seen in Fig. 8. The multiplication of a vector by a number a affects the length and direction of the vector. The product $a\boldsymbol{u}$ stretches the vector $\boldsymbol{u}$ when $a > 1$ and it compresses $\boldsymbol{u}$ when $0 < a < 1$. If $a < 0$ then it reverses the direction of $\boldsymbol{u}$ and it stretches when $a < -1$ and compresses when $-1 < a < 0$. Fig. 8 represents some of these possibilities. Notice that the difference of two vectors is a particular case of the parallelogram law, as it can be seen in Fig. 9.

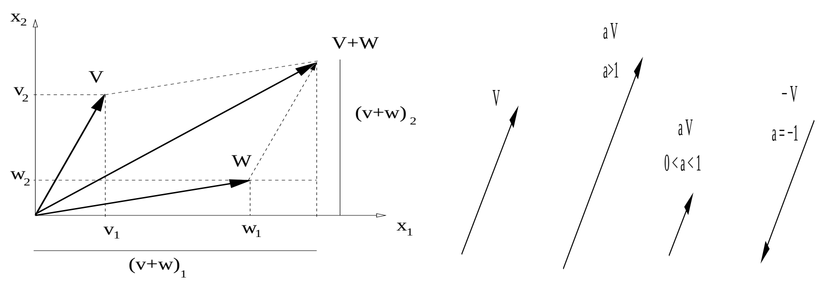

FIGURE 8. The addition of vectors can be computed with the parallelogram law. The multiplication of a vector by a number stretches or compresses the vector, and changes it direction in the case that the number is negative.

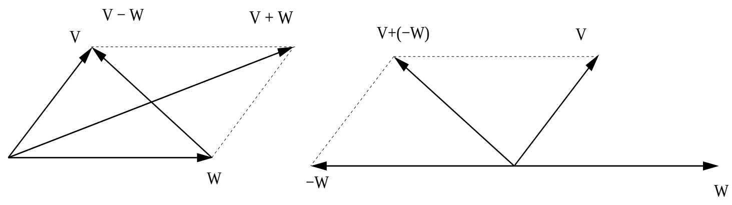

FIGURE 9. The difference of two vectors is a particular case of the parallelogram law of addition of vectors.

The **column picture** interpretation of a general 2×2 linear system given in Eqs. (1.1)-(1.2) is the following: Introduce the coefficient and source column vectors

$$\boldsymbol{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \quad \boldsymbol{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \tag{1.9}$$

and then find the coefficients x_1 and x_2 that change the length of the coefficient column vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ such that they add up to the source column vector $\boldsymbol{b}$, that is,

$$\mathbf{a}_1\, x_1 + \mathbf{a}_2\, x_2 = \mathbf{b}.$$

For example, the column picture of the linear system in Eqs. (1.6)-(1.7) is given in Eq. (1.8). The solution of this system are the numbers $x_1 = 1$ and $x_2 = 2$, and this solution is represented in Fig. 10.

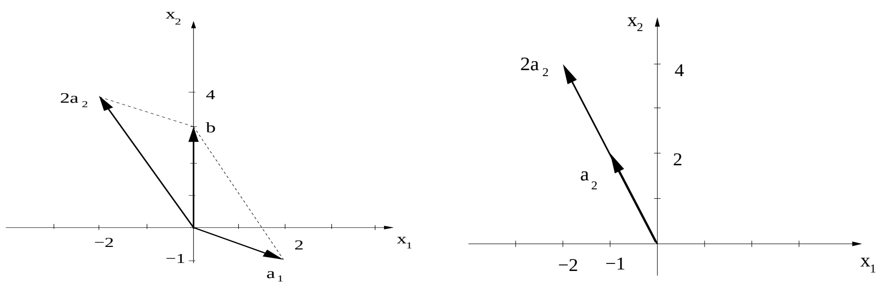

FIGURE 10. Representation of the solution of a 2×2 linear system in the column picture.

The existence and uniqueness of solutions in the case of 2×2 systems can be studied geometrically in the column picture as it was done in the row picture. In this latter case we have seen that all possible 2×2 systems fall into one of these three cases, unique solution, infinitely many solutions and no solutions at all. In Fig. 11 we present these three cases in both row and column pictures.

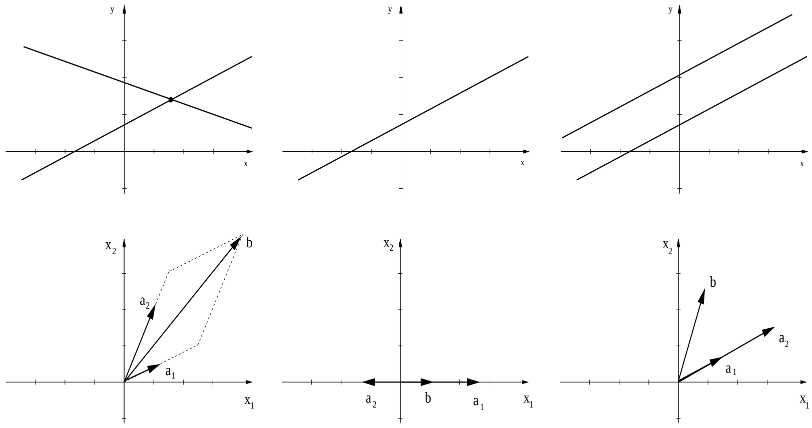

FIGURE 11. Examples of a solutions of general 2×2 linear systems having a unique, infinite many, and no solution, represented in the row picture and in the column picture.

The ideas in the column picture can be generalized to $m \times n$ linear equations, which gives rise to the generalization to m-vectors of the definitions of linear combination presented above. Given m-vectors

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \qquad \boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix},$$

and real numbers a and b, introduce the linear combination of the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ as follows

$$a \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} + b \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} = \begin{bmatrix} au_1 + bv_1 \\ \vdots \\ au_m + bv_m \end{bmatrix}.$$

This definition can be generalized to an arbitrary number of vectors. The m-vector $\boldsymbol{b}$ is a **linear combination** of the m-vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ iff there exist real numbers $x_1, \cdots, x_n$ such that the following equation holds,

$$\boldsymbol{a}_1 x_1 + \cdots + \boldsymbol{a}_n x_n = \boldsymbol{b}.$$

For example, recall the 3×3 system given as the second system in Eqs.(1.5). This system in the column picture is the following: Find numbers x_1, x_2 and x_3 such that

$$\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} x_3 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}. \tag{1.10}$$

These are the main ideas in the column picture. We will see later that linear algebra emerges from the column picture. The next section we give a method, due to Gauss, to solve in an efficient way $m \times n$ linear systems for large m and n.

1.3. **Gauss elimination method.** The Gauss elimination operations (GEO) is a method to find solutions to $m \times n$ linear systems in an efficient way. Efficient means performing as few as possible algebraic steps to find the solution or to show that the solutions does not exist. Before introducing this method, we need several definitions. Consider an $m \times n$ linear system

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

Introduce the matrix of coefficients and the augmented matrix of a linear system, given respectively by the following expressions,

$$A := \overbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}}^{n \text{ columns}} m \text{ rows}, \qquad \big[A|\boldsymbol{b}\big] := \begin{bmatrix} a_{11} & \cdots & a_{1n} & | & b_1 \\ \vdots & & \vdots & | & \vdots \\ a_{m1} & \cdots & a_{mn} & | & b_m \end{bmatrix}$$

We call A an $m \times n$ matrix, and so the augmented matrix of an $m \times n$ linear system is given by the coefficients and the source vector together, so it is an $m \times (n+1)$ matrix. The symbol ":=" denote "definition". For example, in the linear system

$$\begin{aligned} 2x_1 - x_2 &= 0, \\ -x_1 + 2x_2 &= 3, \end{aligned}$$

the matrix of coefficients is 2×2 and the augmented matrix is 2×3, given respectively by

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \qquad \begin{bmatrix} 2 & -1 & | & 0 \\ -1 & 2 & | & 3 \end{bmatrix}.$$

We also use the alternative notation $A = [a_{ij}]$, $\boldsymbol{b} = [b_i]$. Given a matrix $A = [a_{ij}]$, the elements a_{ii} are called **diagonal elements**. Examples of diagonal elements in 3×3, 2×3 and 3×2 matrices are given by the following matrices, where $*$ means a non-diagonal element,

$$\begin{bmatrix} a_{11} & * & * \\ * & a_{22} & * \\ * & * & a_{33} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & * & * \\ * & a_{22} & * \end{bmatrix}, \quad \begin{bmatrix} a_{11} & * \\ * & a_{22} \\ * & * \end{bmatrix}.$$

The **Gauss elimination operations** refers to the following three operations performed on the augmented matrix:

(i) Adding to one row a multiple of the another;
(ii) Interchanging two rows;
(iii) Multiplying a row by a non-zero number.

These operations are respectively represented by the symbols given in Fig. 12.

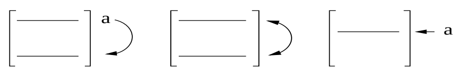

FIGURE 12. A representation of the Gauss elimination operations.

The Gauss elimination operations change the coefficients of the augmented matrix of a system but do not change its solution. Two systems of linear equations having the same solutions are called **equivalent**. It can be shown that there is an algorithm using these operations such that given any $m \times n$ linear system there exists an equivalent system whose augmented matrix is simple in the sense that the solution can be found by inspection. For example, consider the 2×2 linear system in Eq. (1.8), construct its augmented matrix, and perform the following Gauss elimination operations,

$$\begin{bmatrix} 2 & -1 & | & 0 \\ -1 & 2 & | & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & -1 & | & 0 \\ -2 & 4 & | & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & -1 & | & 0 \\ 0 & 3 & | & 6 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 2 & -1 & | & 0 \\ 0 & 1 & | & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 0 & | & 2 \\ 0 & 1 & | & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & | & 1 \\ 0 & 1 & | & 2 \end{bmatrix},$$

and in the last augmented matrix the solution, $x_1 = 1$, $x_2 = 2$ is easy to read. A precise way to define the notion of easy to read is captured in the notion is in echelon form. An $m \times n$ matrix is in **echelon form** iff every element below the diagonal vanishes. Matrices with this property are also called upper triangular. A matrix is in **reduced echelon form** iff it is in echelon form and the first nonzero element in every row satisfies both that it is equal to 1 and it is the only nonzero element in that column. As an example, the following matrices are in echelon form,

$$\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 3 & 2 \\ 0 & 4 & -2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 & 1 \\ 0 & 3 & 4 \\ 0 & 0 & 0 \end{bmatrix}.$$

And the following matrices are not only in echelon form but also in reduced echelon form,

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

Summarizing, the Gauss elimination operations can transform any matrix into a reduce echelon form. Once the augmented matrix of a linear system is written in reduced echelon form, it is not difficult to decide whether the system has solutions or not. For example, suppose that the augmented matrix of a 3×3 linear system has the following reduced echelon form,

$$\begin{bmatrix} 1 & 0 & 2 & | & 1 \\ 0 & 1 & 3 & | & 2 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} x_1 = 1 - 2x_3, \\ x_2 = 2 - 3x_3, \\ x_3 : \text{ free variable.} \end{cases}$$

A variable of an $m \times n$ linear system is called a **free variable** iff for every value of that variable there exists a solution to the linear system. The following result characterizes $n \times n$ linear systems having free variables.

Lemma 1. *An $n \times n$ linear system has k free variables iff the reduced echelon form of its augmented matrix contains k rows of the form* $[0, \cdots, 0|0]$.

We left the proof as an exercise. We remark that the Lemma above is concerned only with linear systems with square matrix of coefficients. The Lemma is not true for $m \times n$ linear systems when $m \neq n$, as the following examples show:

(i) The 3×2 linear system below has one free variable but its associated augmented matrix has no line of the form $[0,0,0|0]$;

$$x_1 + 2x_2 + 3x_3 = 1,$$
$$3x_1 + x_2 + 2x_3 = 2.$$

(ii) The 2×3 linear system below has one line of the form $[0,0|0]$ but it has no free variables;

$$x_1 + 3x_2 = 2,$$
$$2x_1 + 2x_2 = 0,$$
$$3x_1 + x_2 = -2.$$

As an example of the result presented in Lemma 1, consider the 2×2 linear system

$$2x_1 - x_2 = 1 \tag{1.11}$$

$$-\frac{1}{2}\,x_1 + \frac{1}{4}\,x_1 = -\frac{1}{4}. \tag{1.12}$$

It is not difficult to check that Gauss elimination operations can transform the system augmented matrix as follow,

$$\begin{bmatrix} 2 & -1 & | & 1 \\ -\frac{1}{2} & \frac{1}{4} & | & -\frac{1}{4} \end{bmatrix} \to \begin{bmatrix} 2 & -1 & | & 1 \\ 0 & 0 & | & 0 \end{bmatrix},$$

so the system above has a free variable, and therefore, infinitely many solutions.

Lemma 1 presented a condition to determine whether a square linear system has infinity many solutions. We now present a condition on an arbitrary linear system that determines whether the system has no solutions. We first present an example. Consider a linear system with the same matrix coefficients as the one in Eq. (1.11)-(1.12) but with a different source vector:

$$2x_1 - x_2 = 0 \tag{1.13}$$

$$-\frac{1}{2}\,x_1 + \frac{1}{4}\,x_1 = -\frac{1}{4}. \tag{1.14}$$

Multiplying the second equation by -4 one obtains the equation

$$2x_1 - x_2 = 1,$$

whose solutions form a parallel line to the line given in Eq. (1.13). Therefore, the system in Eqs. (1.13)-(1.14) has no solution. Using Gauss elimination operations it is not difficult to check that the system augmented matrix can be transformed as follows,

$$\begin{bmatrix} 2 & -1 & | & 0 \\ -\frac{1}{2} & \frac{1}{4} & | & -\frac{1}{4} \end{bmatrix} \to \begin{bmatrix} 2 & -1 & | & 0 \\ 0 & 0 & | & 1 \end{bmatrix}.$$

The last row has the form $[0,0|1]$, which is a contradiction, therefore the system in Eqs. (1.13)-(1.14) has no solutions. These examples are particular cases of the following result.

Lemma 2. *An $m \times n$ linear system is inconsistent iff the reduced echelon form of its augmented matrix contains a row of the form $[0,\cdots,0|1]$.*

Furthermore, a consistent system contains:

(i) A unique solution iff has no free variables;
(ii) Infinitely many solutions iff it contains at least one free variable.

Proof of Lemma 2: The idea of the proof is to study all possible forms a reduced echelon form can have, one concludes that there are three main cases, no solutions, unique solutions, or infinitely many solutions, according to the form of the reduced echelon for of the system augmented matrix. We only give here the proof in the case of 3×3 linear systems. The reduced echelon form of an inconsistent 3×3 linear system can have only one of the following forms

$$\begin{bmatrix} 1 & 0 & * & | & * \\ 0 & 1 & * & | & * \\ 0 & 0 & 0 & | & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & * & * & | & * \\ 0 & 0 & 0 & | & 1 \\ 0 & 0 & 0 & | & 1 \end{bmatrix};$$

the reduced echelon form of a 3×3 linear system having a the unique solution case has the form

$$\begin{bmatrix} 1 & 0 & 0 & | & * \\ 0 & 1 & 0 & | & * \\ 0 & 0 & 1 & | & * \end{bmatrix};$$

while the reduced echelon form of a 3×3 linear system having infinitely many solutions can have only one of the following the form

$$\begin{bmatrix} 1 & 0 & * & | & * \\ 0 & 1 & * & | & * \\ 0 & 0 & 0 & | & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & * & * & | & * \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix},$$

which correspond to one free variable and two free variables, respectively. □

Exercise: Find all numbers h and k such that the system below has only one, many, or no solutions,

$$x_1 + hx_2 = 1$$
$$x_1 + 2x_2 = k.$$

Solution: Start finding the associated augmented matrix and reducing it into echelon form,

$$\begin{bmatrix} 1 & h & | & 1 \\ 1 & 2 & | & k \end{bmatrix} \to \begin{bmatrix} 1 & h & | & 1 \\ 0 & 2-h & | & k-1 \end{bmatrix}.$$

Suppose $h \neq 2$, for example set $h = 1$, then

$$\begin{bmatrix} 1 & 1 & | & 1 \\ 0 & 1 & | & k-1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & | & 2-k \\ 0 & 1 & | & k-1 \end{bmatrix},$$

so the system has a unique solution for all values of k. (The same conclusion holds if one sets h to any number different of 2.) Suppose now that $h = 2$, then,

$$\begin{bmatrix} 1 & 2 & | & 1 \\ 0 & 0 & | & k-1 \end{bmatrix}.$$

If $k = 1$ then

$$\begin{bmatrix} 1 & 2 & | & 1 \\ 0 & 0 & | & 0 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} x_1 = 1 - 2x_2, \\ x_2 : \text{ free variable.} \end{cases}$$

so there are infinitely many solutions. If $k \neq 1$, then

$$\begin{bmatrix} 1 & 2 & | & 1 \\ 0 & 0 & | & k-1 \neq 0 \end{bmatrix}$$

and the system is inconsistent. Summarizing, for $h \neq 2$ the system has a unique solution for every k. If $h = 2$ and $k = 1$ the system has infinitely many solutions, and if $h = 2$ and $k \neq 1$ the system has no solution. □

1.4. **The span of a set of vectors.** Recall that in Sec. 1.2 we have said that a vector $\boldsymbol{u}$ is a linear combination of the vectors $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ iff there exist real numbers $c_1, \cdots, c_n$ such that

$$\boldsymbol{u} = c_1 \boldsymbol{v}_1 + \cdots + c_n \boldsymbol{v}_n.$$

An important concept in linear algebra is that of Span of a set of vectors. The **Span** of the set of m-vectors $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is the set in $\mathbb{R}^m$ of all possible linear combinations of the vectors $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$. We use the notation

$$\text{Span}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\} \subset \mathbb{R}^m.$$

For example, all possible linear combinations of a single vector $\boldsymbol{v}$ are vectors of the form $c\boldsymbol{v}$, and these vectors belong to a line tangent to $\boldsymbol{v}$. All linear combinations of two vectors $\boldsymbol{v}$, $\boldsymbol{w}$ belong to a plane containing both vectors. Examples of these situations can be seen in Fig. 13.

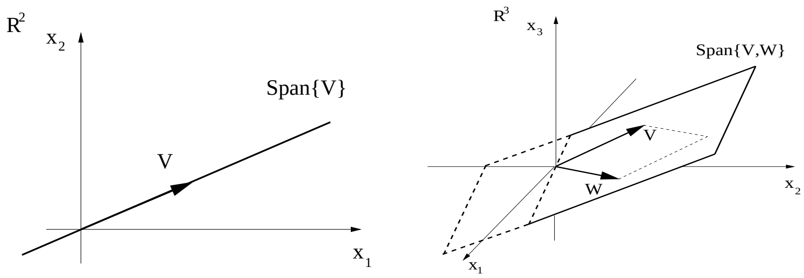

FIGURE 13. Examples of the span of a set of a single vector, and the span of a set of two vectors.

The concept of span enters in the column picture interpretation of a linear system. The $m \times n$ linear system $\boldsymbol{a}_1 x_1 + \cdots + \boldsymbol{a}_n x_n = \boldsymbol{b}$ has a solution iff the source vector $\boldsymbol{b}$ belongs to the Span$\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}$. In Fig. 14 we recall the examples of 2×2 linear systems of the form $\boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 = \boldsymbol{b}$ having a unique solution, infinitely many solutions, and no solution, respectively. In the first two cases the source vector $\boldsymbol{b}$ belongs to Span$\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$, while in the third case $\boldsymbol{b} \notin$ Span$\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$.

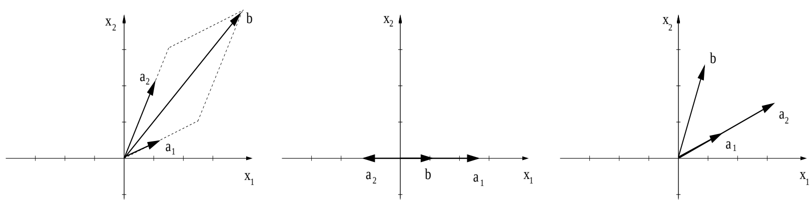

FIGURE 14. Examples of a solutions of general 2×2 linear systems having a unique, infinite many, and no solution, represented in the column picture.

Therefore, in the column picture there are two equivalent interpretations of an $m\times n$ linear system. The first one is: Given the m-vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ and the m-vector $\boldsymbol{b}$, find the real numbers $x_1, \cdots, x_n$ solutions of the equation $\boldsymbol{a}_1\, x_1 + \cdots + \boldsymbol{a}_n\, x_n = \boldsymbol{b}$. The second interpretation is then the following: Given the m-vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ and the m-vector $\boldsymbol{b}$, find the real numbers $x_1, \cdots, x_n$ such that $\boldsymbol{b} \in \mathrm{Span}\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}$.

1.5. **Matrix-vector product.** Matrices have been introduced in Sec. 1.3 as a way to simplify the calculations needed to find solutions of linear systems. For example, the augmented matrix of a linear system is an object containing all the equation coefficients and the equations sources which are used in the Gauss elimination operations in order to obtain the solution to that linear system. In this section we express a linear system in yet another way, introducing an operation between matrices and vectors. Consider an $m \times n$ linear system given by

$$\boldsymbol{a}_1 x_1 + \cdots + \boldsymbol{a}_n x_n = \boldsymbol{b},$$

and introduce the matrix of coefficients $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$. The new idea is to consider the numbers $x_1, \cdots, x_n$ as an n-vector $\boldsymbol{x} \in \mathbb{R}^n$ given by

$$\boldsymbol{x} := \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

The **matrix-vector** product of the $m \times n$ matrix A and the n-vector $\boldsymbol{x}$ given above is defined as follows,

$$A\boldsymbol{x} := [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} := \boldsymbol{a}_1 x_1 + \cdots + \boldsymbol{a}_n x_n.$$

Therefore, an $m \times n$ linear system can be presented in the following way: Given an $m \times n$ matrix A and an m-vector $\boldsymbol{b}$, find an n-vector $\boldsymbol{x}$ solution of the matrix-vector product equation $A\boldsymbol{x} = \boldsymbol{b}$. As an example, consider the 2×2 linear system given in Eq. (1.8). Introduce then the vector

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Then the linear system above can be written as

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Let us consider a second example, given by the 2×3 linear system in the unknowns x_1 and x_2 given by

$$\begin{aligned} x_1 - x_2 &= 0, \\ -x_1 + x_2 &= 2, \\ x_1 + x_2 &= 0. \end{aligned}$$

The column picture interpretation is

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} x_1 + \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} x_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}.$$

while the matrix-vector product interpretation is

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}.$$

Consider a third example, the 2×3 linear system

$$\begin{aligned} 2x_1 - x_2 + x_3 &= 0, \\ -x_1 + 2x_2 - x_3 &= 3. \end{aligned}$$

The column picture interpretation is

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix} x_1 + \begin{bmatrix} -1 \\ 2 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ -1 \end{bmatrix} x_3 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

while the matrix-vector product interpretation is

$$\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

The matrix-vector product introduced above satisfies several properties, and one of those properties we are most interested in is that the matrix-vector product preserves linear combination. This property is summarized in the following result.

Theorem 2. *If A is an $m \times n$ matrix, $\boldsymbol{u}$, $\boldsymbol{v}$ are arbitrary n-vectors, and a, b are arbitrary real numbers, then the matrix-vector product satisfies the following equation*

$$A(a\boldsymbol{u} + b\boldsymbol{v}) = aA\boldsymbol{u} + bA\boldsymbol{v}.$$

This Theorem says that the matrix-vector product of a linear combination of vectors is the linear combination of the matrix-vector products.

Proof of Theorem 2: This property follows directly from the definition of the matrix-vector product:

$$\begin{aligned} A(a\boldsymbol{u} + b\boldsymbol{v}) &= [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n] \begin{bmatrix} au_1 + bv_2 \\ \vdots \\ au_n + bv_n \end{bmatrix} \\ &= \boldsymbol{a}_1(au_1 + bv_2) + \cdots + \boldsymbol{a}_n(au_n + bv_n) \\ &= a\big(\boldsymbol{a}_1 u_1 + \cdots + \boldsymbol{a}_n u_n\big) + b\big(\boldsymbol{a}_1 v_1 + \cdots + \boldsymbol{a}_n v_n\big) \\ &= aA\boldsymbol{u} + bA\boldsymbol{v}. \end{aligned}$$

□

As usual, the expression above contains the particular cases $a = b = 1$ and $b = 0$, which are given, respectively, by

$$A(\boldsymbol{u} + \boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v}, \qquad A(a\boldsymbol{u}) = aA\boldsymbol{u}.$$

1.6. **Homogeneous linear systems and Spans.** An $m \times n$ linear system $A\boldsymbol{x} = \boldsymbol{b}$ is called **homogeneous** iff the source vector $\boldsymbol{b} = \boldsymbol{0}$. This type of linear systems possess several interesting properties, the simplest one is that they always have at least one solution, $\boldsymbol{x} = \boldsymbol{0}$, called the trivial solution. An equally important property is that homogeneous linear systems can also have non-zero solutions. Consider the following example: Find the solutions $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ of the system

$$\begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{1.15}$$

The solution of this system can be found using Gauss elimination operations, as follows

$$\begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} \to \begin{bmatrix} 2 & -4 \\ 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} x_1 = 2x_2, \\ x_2 \text{ free variable.} \end{cases}$$

Therefore, the set of all solutions of the linear system above is given by

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} x_2, \quad \forall x_2 \in \mathbb{R} \quad \Rightarrow \quad \boldsymbol{x} \in \operatorname{Span}\Big\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix} \Big\}.$$

The expression above means that all solutions to the linear system given in Eq. (1.15) are elements in the span of the vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Therefore, the set of all solutions of this linear system can be identified with the set of points that belong to the line shown in Fig. 15.

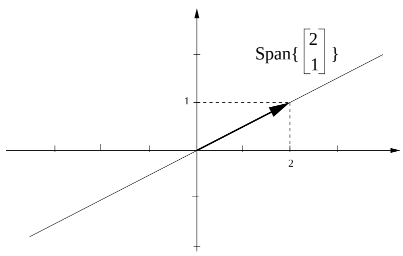

FIGURE 15. Solutions of the homogeneous linear system given in Eq. (1.15).

From the example above we find out that there exists a relation between non-trivial solutions to homogeneous linear systems and free variables of this system. This relation is summarized in the following result.

Proposition 1. *An $m \times n$ homogeneous linear system has non-trivial solutions iff the system has free variables.*

Proof of Proposition 1: Suppose that the vector $\boldsymbol{x}$ is the solution to an homogeneous linear system, and suppose that the component x_1 of the vector $\boldsymbol{x}$ is a free variable. Then, the component x_1 can take any value and the vector $\boldsymbol{x}$ is a solution of the linear system, so then take $x_1 = 1$, which then implies that there exists a non-zero vector $\boldsymbol{x}$ solution of the linear system. $\square$

The non-trivial solutions of homogeneous $m \times n$ linear systems $A\boldsymbol{x} = \boldsymbol{0}$, where the coefficient matrix $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ is formed by column vectors $\boldsymbol{a}_i \in \mathbb{R}^m$, $i = 1, \cdots m$, can be expressed in terms of spans of vectors in $\mathbb{R}^n$. As an example, consider the following 2×3 linear system,

$$\begin{bmatrix} 2 & -2 & 4 \\ 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{1.16}$$

The coefficient matrix has column vectors that belong to $\mathbb{R}^2$. Let us now compute the solutions to the homogeneous linear system above. We start with Gauss elimination operations

$$\begin{bmatrix} 1 & 3 & 2 \\ 2 & -2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 2 \\ 0 & -8 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} x_1 = -2x_3, \\ x_2 = 0, \\ x_3 \text{ free variable.} \end{cases}$$

Therefore, the set of all solutions of the linear system above is given by

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_3 \\ 0 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} x_3, \quad \forall x_3 \in \mathbb{R} \quad \Rightarrow \quad \boldsymbol{x} \in \mathrm{Span}\Big\{ \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \Big\}.$$

In Fig. 16 we emphasize that the vectors $\boldsymbol{x}$, which are solutions of the homogeneous linear system above, belong to the space $\mathbb{R}^3$, while the column vectors of the coefficient matrix of this same system belong to the space $\mathbb{R}^2$.

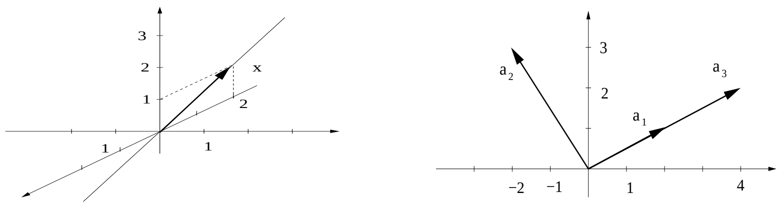

FIGURE 16. The picture on the left represents the solutions of the homogeneous linear system given in Eq. (1.16), which are elements in the space $\mathbb{R}^3$. The picture on the right represents the column vectors of the coefficient matrix in this system given in Eq. (1.16), which are vectors in the space $\mathbb{R}^2$.

Knowing the solutions of an homogeneous linear system gives information about the solutions of an inhomogeneous linear system with the same coefficient matrix. The next result establishes this relation in a precise way.

Theorem 3. *If the $m \times n$ linear system $A\boldsymbol{x} = \boldsymbol{b}$ is consistent, and the vector $\boldsymbol{x}_p$ is one particular solution of this linear system, then all solutions to this linear system are vectors of the form $\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_h$, where the vector $\boldsymbol{x}_h$ is any solution of the homogeneous linear system $A\boldsymbol{x}_h = 0$.*

Proof of Theorem 3: Let the vector $\boldsymbol{x}$ be any solution of the linear system above, that is, $A\boldsymbol{x} = \boldsymbol{b}$. Given the particular solution $\boldsymbol{x}_p$ of this linear system, introduce the vector $\hat{\boldsymbol{x}} := \boldsymbol{x} - \boldsymbol{x}_p$. Then, this vector $\hat{\boldsymbol{x}}$ satisfies the equation

$$A\hat{\boldsymbol{x}} = A(\boldsymbol{x} - \boldsymbol{x}_p) = A\boldsymbol{x} - A\boldsymbol{x}_p = \boldsymbol{b} - \boldsymbol{b} = \mathbf{0}.$$

Therefore, the vector $\hat{\boldsymbol{x}}$ is solution of the homogeneous linear system $A\hat{\boldsymbol{x}} = \mathbf{0}$. This establishes the Theorem. □

We say that the solutions of a linear system are expressed in **parametric form** iff the solution vector $\boldsymbol{x}$ is written as is described in Theorem 3, that is, $\boldsymbol{x} = \boldsymbol{x}_p + \boldsymbol{x}_h$, where the vector $\mathbf{x}_h$ is a solution of the homogeneous linear system, and the vector $\boldsymbol{x}_p$ is any solution of the inhomogeneous linear system.

Exercise: Find all solutions of the 2×2 linear system below, and write them in parametric form, where the linear system is given by

$$\begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}. \tag{1.17}$$

Solution: Find the solutions of this inhomogeneous linear system using Gauss elimination operations,

$$\begin{bmatrix} 1 & -2 & | & 3 \\ 2 & -4 & | & 6 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & | & 3 \\ 0 & 0 & | & 0 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} x_1 = 2x_2 + 3, \\ x_2 \text{ free variable.} \end{cases}$$

Therefore, the set of all solutions of the linear system above is given by

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_2 + 3 \\ x_2 \end{bmatrix} \quad \Rightarrow \quad \boxed{\boldsymbol{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} x_2 + \begin{bmatrix} 3 \\ 0 \end{bmatrix}}.$$

□

The solution in the exercise above can also be expressed in the form

$$\boldsymbol{x} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} + \boldsymbol{x}_h, \qquad \boldsymbol{x}_h \in \mathrm{Span}\Big\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\Big\}.$$

In Fig. 17 we represent these solutions on the plane.

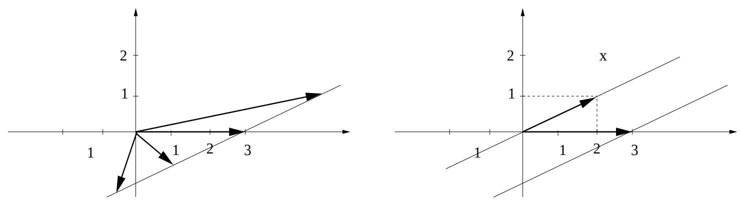

FIGURE 17. The picture on the left represents four solutions of the inhomogeneous linear system given in Eq. (1.17), which are vectors ending in the line drawn in that picture. The picture on the right represents the line associated with the space $\mathrm{Span}\{\left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]\}$, which pass through the origin, and the line associated with the solutions of the inhomogeneous system given in Eq. (1.17).

1.7. **Linear dependence and independence.** A non-empty set of m-vectors $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is **linearly independent** (l.i.) iff the $m \times n$ homogeneous linear system $V\boldsymbol{x} = \boldsymbol{0}$ has only the trivial solution $\boldsymbol{x} = \boldsymbol{0}$, where the matrix V is given by $V = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]$ and the the vector $\boldsymbol{x} \in \mathbb{R}^n$. The set of vectors above is called **linearly dependent** (L.d.) iff there exists a non-trivial solution, $\boldsymbol{x} \neq \boldsymbol{0}$, of the linear system $V\boldsymbol{x} = \boldsymbol{0}$.

Appendix A. Notation and conventions

Vectors will be denoted by boldface letters, like $\boldsymbol{a}$ and $\boldsymbol{b}$. Matrices will be denoted by capital letters like A and B. We list below several mathematical symbols used in these notes:

$:=$	Definition,	$\Longrightarrow$	Implies,
$\forall$	For all,	$\exists$	exists,
Proof.	Begining of a proof,	$\square$	End of a proof.