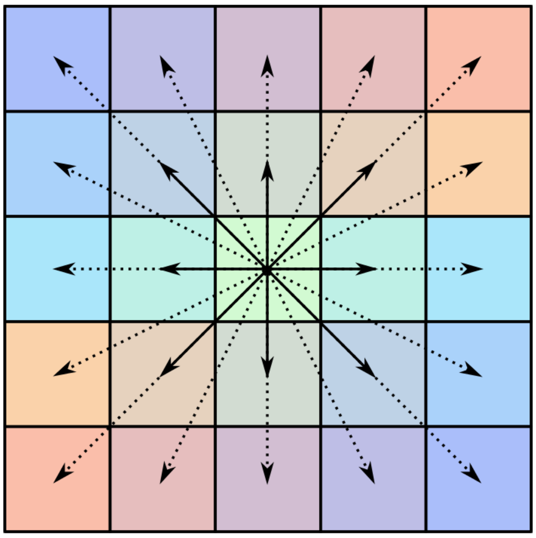
Gymnázium Evolution Jižní Město



Intro-ish To Linear Algebra

Madam Pekáč

November 7, 2024

Preface

This text covers selected topics from the curriculum of a typical undergraduate linear algebra course. Almost no pre-existing knowledge is strictly required save a superficial understanding of propositional logic and set theory. A reasonably good ability to manipulate algebraic expressions should prove advantageous, too.

Mathematics is an exact and rigorous language. Words and symbols have singular, precisely defined, meaning. Many students fail to grasp that intuition and imagination are paramount, but they serve as a *starting point*, with formal logical expression being the end. For example, an intuitive understanding of a *line* as an infinite flat 1D object is pretty much correct but not *formal.* It is indeed the formality of mathematics which puts many students off. Whereas high school mathematics is mostly algorithmic and non-argumentative, higher level maths tends to be the exact opposite – full of concepts and relations between those, which one is expected to be capable of grasping and formally describing. Owing to this, I wish this text would be a kind of synthesis of the formal and the conceptual. On one hand, rigorous definitions and proofs are given; on the other, illustrations, examples and applications serve as hopefully efficient conveyors of the former's geometric nature.

Linear algebra is a mathematical discipline which studies – as its name rightly suggests – the *linear.* Nevertheless, the word *linear* (as in 'line-like') is slightly misplaced. The correct term would perhaps be *flat* or, nigh equivalently, *not curved.* It isn't hard to imagine why curved objects (as in *geometric* objects, say) are more difficult to describe and manipulate than objects flat. For instance, the formula for the volume of a cube is just the product of the lengths of its sides. Contrast this with the volume of a still 'simple', yet curved, object – the ball. Its volume cannot even be *precisely* determined; its calculation involves approximating an irrational constant and the derivation of its formula is starkly unintuitive without basic knowledge of measure theory.

As such, linear algebra is a highly 'geometric' discipline and opportunities for visual interpretations abound. This is also a drawback in a certain sense. One should not dwell on visualisations alone as they tend to lead astray where imagination falls short. Symbolic representation of the geometry at hand is key.

The word *linear* however dons a broader sense in modern mathematics. It can be rephrased as reading, 'related by addition and multiplication by a scalar'. We trust kind readers have been acquainted with the notion of a *linear function.* A linear function is (rightly) called *linear* for it receives a number as input and outputs its *constant* multiple plus another *constant* number. Therefore, the output is in a *linear* relation to the input – it is multiplied by some fixed number and added to another. This understanding of the word is going to prove crucial already in the first chapter, where we study *linear systems.* Following are *vector spaces* and *linear maps*, concepts whose depth shall occupy the span of this text. Each chapter is further endowed with an *applications* section where I try to draw a simile between mathematics and common sense.

Contents

Chapter 1

Linear Systems

Linear systems are by definition sets of linear equations, that is, of equations which relate present variables in a *linear* way. It is important to understand what this means. Spelled out, an expression on either side of any of the equations is formed *solely* by

(1) multiplying the variables by a given number (**not another variable**),

(2) adding these multiples together.

Any such combination where variables are only allowed to be multiplied by a constant and added is called a *linear combination*. This term is extremely important and ubiquitous throughout the text; hence, it warrants an isolated definition.

Definition 1.0.1 (Linear combination)

Let $x_1, \ldots, x_n$ with $n \in \mathbb{N}$ be variables. Their *linear combination* is any expression of the form

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n,$$

where $a_1, \ldots, a_n$ are numbers.

Remark 1.0.2

In the definition above, we have deliberately not specified what type of *numbers* we mean. In the future, we shall work extensively with real and complex numbers as well as elements of other fields, which dear readers might not have even recognised as 'numbers' thus far. The only important concept in this regard is the clear distinction between a *number* (later *scalar*) and a *variable* (later *vector*).

Example 1.0.3

Consider the variables x, y and z. The expression

$$3x + 2y - 0.5z$$

is their linear combination whereas

$$5x + 3y - yz + 7z^2$$

is not.

To reiterate, a *linear system* is any set of equations featuring only linear combinations of variables; these equations are consequently called *linear* as well. A *solution* of a linear system is the set of all possible substitutions of numbers (in place of variables) which make the equations true.

It is clear that every linear equation can be rearranged to

$$a_1x_1 + \cdots + a_nx_n = c$$

for some variables $x_1, \ldots, x_n$ and numbers $a_1, \ldots, a_n, c$ by simple subtraction. This is how we shall define it, for simplicity.

Definition 1.0.4 (Linear equation)

Any equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = c, \tag{1.1}$$

where $x_1, \ldots, x_n$ are variables and $a_1, \ldots, a_n, c$ are numbers, is called *linear*. A *solution* of a linear equation is an n-tuple $(b_1, \ldots, b_n)$ of numbers such that under the substitutions $x_i \coloneqq b_i$, for $i \in \{1, \ldots, n\}$, the equation (1.1) is satisfied.

Example 1.0.5

The equation

$$3x_1 - 2x_2 + 4x_3 + x_4 = 5$$

is linear in variables x_1, x_2, x_3 and x_4. On the contrary,

$$3x_1x_2 - 4x_3^2 = 10$$

is **not** linear.

Definition 1.0.6 (Linear system)

Any set of linear equations in the given variables $x_1, \ldots, x_n$ is called a *linear system*. A *solution* of a linear system is an n-tuple $(b_1, \ldots, b_n)$ which solves every linear equation in the set.

Example 1.0.7

The set of equations

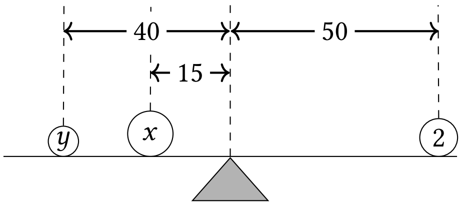
$$\begin{array}{rcrcrcr} 3x_1 & - & x_2 & + & 2x_3 & = & 1 \\ x_1 & & & - & x_3 & = & -1 \\ 2x_1 & - & 3x_2 & + & 3x_3 & = & 0 \end{array}$$

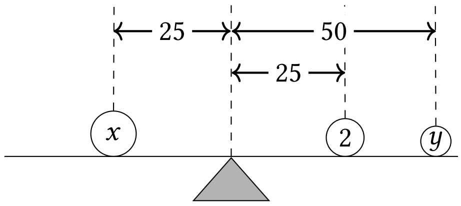
is a linear system whose solution is the triple $(0, 1, 1)$.

We proceed to discuss two trivial examples, which readers might have discussed in high school, naturally leading to linear systems. More sophisticated examples are presented in the applications section.

Example 1.0.8 (Static equations)

Suppose we have three objects – one with a mass of 2 and the other two with masses unknown. Experimentation produces these two balances.





For the weights to be in balance, the sums of *moments* on both sides of the scales must be identical one to another. A *moment* of an object is its distance from the centre of the scales times its mass. This condition yields a system of two linear equations

$$\begin{aligned} 15x + 40y &= 50 \cdot 2, \\ 25x &= 25 \cdot 2 + 50y. \end{aligned}$$

Or, after rearrangement (to stay true to our definition of linear equation),

$$\begin{aligned} 15x + 40y &= 50 \cdot 2, \\ 25x - 50y &= 25 \cdot 2. \end{aligned}$$

Example 1.0.9 (Chemical reactions)

Toluene, $\mathtt{C_7H_8}$, mixes (under right conditions) with nitric acid, $\mathtt{HNO_3}$, to produce trinitrotoluene (widely known as TNT), $\mathtt{C_7H_5O_6N_3}$, along with dihydrogen monoxide, $\mathtt{H_2O}$. If we want this chemical reaction to occur successfully, we must (among other things) ascertain we mix the constituents in the right proportion. In pseudo-chemical notation, the reaction to take place can be written as

$$x \cdot \mathtt{C_7H_8} + y \cdot \mathtt{HNO_3} \longrightarrow z \cdot \mathtt{C_7H_5O_6N_3} + w \cdot \mathtt{H_2O}.$$

Comparing the number of atoms of each element before the reaction and afterwards (which must remain identical owing to the conservation of energy) yields the system

$$\begin{aligned} \mathtt{H}: 8x + 1y &= 5z + 2w, \\ \mathtt{C}: 7x \quad\;\; &= 7z, \\ \mathtt{N}: \quad\;\; 1y &= 3z, \\ \mathtt{O}: \quad\;\; 3y &= 6z + 1w. \end{aligned}$$

In the next section, we devise an algorithm to solve any system of linear equations.

1.1 Describing Solution Sets of Linear Systems

In section **??**, we studied specific (simple) classes of linear systems and touched upon a few important concepts, including, but not limited to, *parameters*, *free variables*, *underdetermined* and *overdetermined* systems.

We continue down this road and bring a general description of solution sets of linear systems. Before we formulate the result we shall endeavour to prove in this section, we introduce a few pieces of notation which are going to allow us to manipulate linear systems more efficiently. Do note that behind these mere 'pieces of notation' there lies hidden a much deeper geometric meaning, to be uncovered in later chapters.

Definition 1.1.1 (Matrix)

An $m \times n$ *matrix* is an array of numbers with m rows and n columns. The numbers are then called *entries* of the matrix.

Matrices allow us to write linear systems in a much more succinct manner. For example, the system

$$\begin{aligned} -x + y &= 2 \\ 2x - 2y &= 5 \end{aligned}$$

can be written using a matrix like this:

$$\left(\begin{array}{cc|c} -1 & 1 & 2 \\ 2 & -2 & 5 \end{array}\right),$$

abusing the fact that the same variables are piled in a single column and each row is a single linear equation. The bar on the right side simply serves to divide left sides of the equations from right ones.

Matrices make (amongst other things) Gauss-Jordan elimination easier to perform and keep track of its progress. The matrix of the eliminated system looks like this

$$\left(\begin{array}{cc|c} -1 & 1 & 2 \\ 0 & 0 & 9 \end{array}\right)$$

and has been reached by the row operation $\mathtt{II} + 2\mathtt{I}$.

Certain matrices are special (for reasons soon to be revealed) and we call them *vectors*.

Definition 1.1.2 (Vector)

A *column vector* is an $n \times 1$ matrix (that is, matrix with a single column) and a *row vector* is a $1 \times n$ matrix (a matrix with a single row). As column vectors are the 'default', we call them simply *vectors*.

There exists an obvious bijection between tuples $(v_1, \dots, v_n)$ and column vectors $\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$. Consequently, we say that a vector $\mathbf{v}$ with entries $v_1, \dots, v_n$ *solves* a linear equation

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = c$$

if the tuple $(v_1, \ldots, v_n)$ does.

The addition of vectors and their multiplication by a number are defined naturally.

Definition 1.1.3 (Adding vectors)

Given vectors

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \quad \text{and} \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix},$$

their *sum* is defined as the vector

$$\mathbf{u} + \mathbf{v} := \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix}.$$

Definition 1.1.4 (Multiplying vector by a number)

Given a vector

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

and a number c, the *scalar c-multiple of* $\mathbf{v}$ is the vector

$$c\mathbf{v} := \begin{pmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{pmatrix}.$$

The multiplying number c is often referred to as a *scalar*.

We further need to discuss the concept of *free variables* and *parameters*.

In the previous section, we described the solution set of the system (**??**) using three different ways. In each case, two of the variables were independent and the third was their linear combination. We style the two independent variables, *parameters*. Vaguely said, a *parameter* is a variable on the value thereof other variables depend.

The question arises: 'Which variables to choose as *parameters*?' The answer descends: 'Why, of course, my child, choose the *free variables*!' After the process of Gauss-Jordan elimination, a preceding row always has more variables present than its neighbour downstairs. Occasionally, the number of additional variables is larger than one. It is clear that in such cases, back-substitution cannot determine the values of those additional variables exactly (as it leads to a linear equation in more than one variable). All save one of those variables are to be chosen as *parameters* and serve the noble purpose of describing the value of the last variable standing. Custom dictates that all but the leftmost variable in such a row are labelled *free* and the leftmost variable called a *pivot*. In light of this, the heavenly answer can be decrypted – the *free* variables shall serve as *parameters* and the value of the *pivot* written as a linear combination of free variables.

To understand explicitly the preceding paragraph, consider the eliminated system

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 3 & 1 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Row `II` has two more variables than row `III` and row `I` also wins by two variables over row `II`. As the holy text states, the variable x_4 of the fourth column is *free*, whereas x_3 is a *pivot*. Therefore, x_4 now serves as a parameter and from row `II` we get the relation

$$x_3 = -x_4 + 4.$$

Row `I` brings in a new free variable – x_2 – and a new pivot – x_1. Using the fact that x_3, the pivot from row `II`, is already expressed as a linear combination of free variables, we substitute into row `I` to get

$$x_1 + 2x_2 - (-x_4 + 4) + 3x_4 = 1.$$

A tiny bit of cheap computation yields

$$x_1 = -2x_2 - 4x_4 + 5.$$

Hence, all the pivots of the systems are expressed as linear combinations of free variables. The set of solutions of this system can be described as the set of quadruples $(-2x_2 - 4x_4 + 5, x_2, -x_4 + 4, x_4)$.

Visualisation of the concepts of pivots and free variables is provided in figure 1.2.

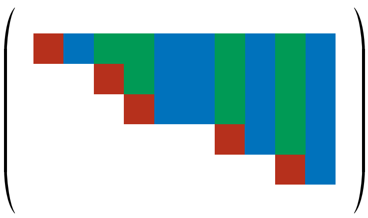


Figure 1.2: Visual depiction of an eliminated matrix. Red variables are pivots, blue ones are free and green ones are pivots from lower rows.

Using vectors, the solution set of the currently studied system can be expressed quite elegantly. First, the quadruple $(-2x_2 - 4x_4 + 5, x_2, -x_4 + 4, x_4)$ corresponds to the column vector

$$\begin{pmatrix} -2x_2 - 4x_4 + 5 \\ x_2 \\ -x_4 + 4 \\ x_4 \end{pmatrix}.$$

This vector can be further broken down into three vectors, two for the free variables and one for the constants. Explicitly,

$$\begin{pmatrix} -2x_2 - 4x_4 + 5 \\ x_2 \\ -x_4 + 4 \\ x_4 \end{pmatrix} = x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -4 \\ 0 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 5 \\ 0 \\ 4 \\ 0 \end{pmatrix}.$$

Take note that the last vector is a *particular* solution of the system obtained by setting $x_2 = x_4 = 0$. Adding random multiples of the vectors

$$\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -4 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

to this particular solution generates more solutions of the system.

Let's make another example, shall we? In this eliminated system of two equations in three variables,

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 2 \\ 0 & 0 & 2 & 2 \end{array}\right),$$

the variables x_1 and x_3 are pivots and x_2 is free. Judging from the previous example, we should be able to express its solution as $\mathbf{u} + x_2\mathbf{v}$ where $\mathbf{u}$ and $\mathbf{v}$ are vectors and, furthermore, $\mathbf{u}$ is some particular solution of the system at hand.

Indeed, choosing x_2 to be a parameter, back-substitution yields $x_3 = 1$ and $x_1 = 2-x_2+x_3 = -x_2+3$. Hence, every vector of the shape

$$\begin{pmatrix} -x_2 + 3 \\ x_2 \\ 1 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$$

solves the system.

We're now equipped to formulate a result about the 'shape' of a linear system's solution set with a rather far-reaching importance.

Theorem 1.1.5 (Solution set of a linear system)

The solution set of every linear system can be written in the form

$$\{\mathbf{u} + t_1\mathbf{v_1} + t_2\mathbf{v_2} + \ldots + t_l\mathbf{v_l}\},$$

where $\mathbf{u}$ is a particular solution, $\mathbf{v_1}, \ldots, \mathbf{v_l}$ are vectors and $t_1, \ldots, t_l$ are parameters corresponding to the free variables of the eliminated system.

Before the proof, we formulate an immediate corollary.

Corollary 1.1.6 (Number of solutions of a linear system)

Every linear system has zero, one or infinitely many solutions.

Proof. Referring to the form of the solution set of a linear system from theorem 1.1.5, we distinguish three cases:

(1) The vector $\mathbf{u}$ doesn't exist, therefore the system has *no solution.*

(2) The vector $\mathbf{u}$ exists and there are no free variables (only pivots) in the eliminated system. In this case, the solution is *unique.*

(3) The vector $\mathbf{u}$ exists and there is at least one free variable to be found in the eliminated system. In this case, the substitution of any number in place of the free variables generates a solution. Hence, there are *infinitely many.* ■

On our way to the proof of theorem 1.1.5, we make a preparatory step. We call a linear system *homogeneous* if the right side of its every equation is 0. Concretely, a *homogeneous* linear system assumes the form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & 0 \\ & & & & & & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & 0. \end{array}$$

Notice that this system always has at least one solution, namely the vector $\mathbf{0}$ – the vector whose every entry is 0. We shall first prove the following proposition.

Proposition 1.1.7 (Solution set of a homogeneous linear system)

The solution set of a homogeneous *linear system can be written in the form*

$$\{t_1\mathbf{v_1} + t_2\mathbf{v_2} + \dots + t_l\mathbf{v_l}\},$$

where $\mathbf{v_1}, \dots, \mathbf{v_l}$ *are vectors and* $t_1, \dots, t_l$ *are parameters corresponding to the free variables of the eliminated system.*

Proof. We consider a homogeneous linear system as above:

$$\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,n}x_n & = & 0 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \dots & + & a_{2,n}x_n & = & 0 \\ & & & & & & & \vdots & \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \dots & + & a_{m,n}x_n & = & 0. \end{array} \tag{1.2}$$

Firstly, in the light of theorem **??**, we may assume that the system has been reduced to echelon form. We shall prove that every pivot can be written as a linear combination of free variables by induction on the number k of rows (counting from the bottom) already substituted into. This basically mimics the traditional back-substitution process.

Without loss of generality, we may also assume that no rows full of zeroes are left at the bottom of the system, as those can be ignored. Hence, the last row of the eliminated linear system looks like this:

$$a_{m,j}x_j + a_{m,j+1}x_{j+1} + \dots + a_{m,n}x_n = 0$$

for adequate $1 \leq j \leq n$ and $a_{m,j} \neq 0$. Here, x_j is the pivot and $x_{j+1}, \dots, x_n$ are free. This gives the expression

$$x_j = -\frac{1}{a_{m,j}}(a_{m,j+1}x_{j+1} + \dots + a_{m,n}x_n)$$

of the pivot x_j as a linear combination of the free variables $x_{j+1}, \dots, x_n$. So, the result holds for $k = 0$.

Now, supposing all pivots in the last k rows of the system (1.2) have been written as linear combinations of free variables, we write the pivot of the $(m-k)$-th row (or $(k+1)$-st from the bottom) also as a linear combination of free variables. Again, there exists some smallest $1 \leq i \leq n$ such that $a_{m-k,i} \neq 0$. The $(m-k)$-th row is thus

$$a_{m-k,i}x_i + a_{m-k,i+1}x_{i+1} + \ldots + a_{m-k,n}x_n = 0.$$

Performing an analogous computation gives

$$x_i = -\frac{1}{a_{m-k}}(a_{m-k,i+1}x_{i+1} + \ldots + a_{m-k,n}x_n). \tag{1.3}$$

All the variables found on the right side of (1.3) are either free or pivots from lower rows. However, by induction hypothesis, all pivots from lower rows have already been expressed as linear combinations of free variables. Simple substitution now yields an expression of x_i as a linear combination of free variables. With l denoting the number of free variables of the eliminated system and splitting the solution vector into a sum of scalar multiples of free variables, the result is proven. ■