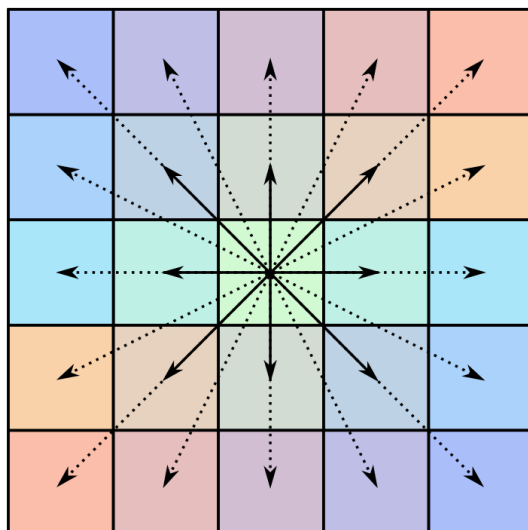


GYMNÁZIUM EVOLUTION JIŽNÍ MĚSTO



Intro-ish To Linear Algebra

Madam Pekáč

September 15, 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 when 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

1	Linear Systems	7
1.1	Gauss-Jordan Elimination	10

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

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

where a_1, \dots, 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, \dots, x_n and numbers a_1, \dots, 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

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = c \quad (1.1)$$

where x_1, \dots, x_n are variables and a_1, \dots, a_n, c are numbers, is called *linear*. A *solution* of a linear equation is an n -tuple (b_1, \dots, b_n) of numbers such that under the substitutions $x_i := b_i$, for $i \in \{1, \dots, n\}$, 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, \dots, x_n is called a *linear system*. A *solution* of a linear system is an n -tuple (b_1, \dots, b_n) which solves every linear equation in the set.

Example 1.0.7

The set of equations

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

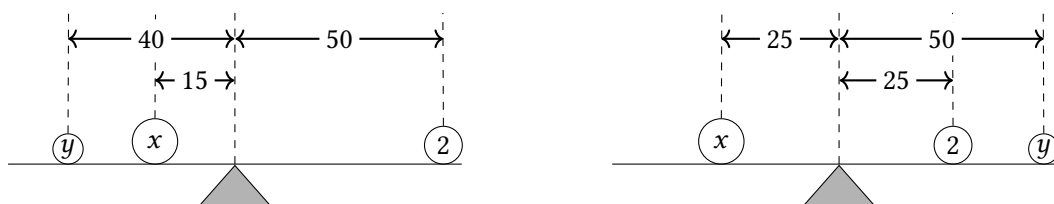
is a *linear system* whose solution is the triple $(0, 1, 1)$.

We proceed to discuss two trivial examples, which readers might have discussed in high school, that naturally lead 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 sum of *moments* on either side of the scales must be identical. 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

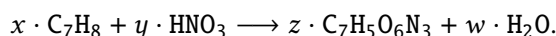
$$\begin{aligned} 15x + 40y &= 50 \cdot 2, \\ 25x &= 25 \cdot 2 + 50y. \end{aligned}$$

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, C_7H_8 , mixes (under right conditions) with nitric acid, HNO_3 , to produce trinitrotoluene (widely known as TNT), $C_7H_5O_6N_3$, along with dihydrogen monoxide, 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



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

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

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

1.1 Gauss-Jordan Elimination

Probably the most well-known algorithm for solving a [linear system](#) is the *Gauss-Jordan Elimination*. As its name partially implies, its heart lies in the *elimination* of variables one by one, until only a single linear equation in one variable stands unsolved. This is done by applying different *transformations* to the initial system that are guaranteed not to alter the solution. We're going to solve a linear system first and describe the general method second.

Problem 1.1.1

Solve the linear system

$$\begin{aligned} 3x_3 &= 9 \\ x_1 + 5x_2 - 2x_3 &= 2. \\ \frac{1}{3}x_1 + 2x_2 &= 3 \end{aligned}$$

SOLUTION. We aim to transform the system step by step to a form which allows us to (successively) eliminate all variables.

The first transformation entails a simple exchange of the first and third row.

$$\text{Swapped first and third row.} \quad \begin{array}{rcl} \frac{1}{3}x_1 + 2x_2 & = & 3 \\ x_1 + 5x_2 - 2x_3 & = & 2. \\ 3x_3 & = & 9 \end{array}$$

Next, we shall scale the first row by a factor of 3.

$$\text{Scaled the first row by 3.} \quad \xrightarrow{\cdot 3} \begin{array}{rcl} x_1 + 6x_2 & = & 9 \\ x_1 + 5x_2 - 2x_3 & = & 2. \\ 3x_3 & = & 9 \end{array}$$

Finally, we subtract the first row from the second row. Said in a more foreshadowing manner, we add the (-1) -multiple of the first row to the second row.

$$\text{Subtracted the first row from the second.} \quad \cdot(-1) \quad \begin{array}{rcl} x_1 + 6x_2 & = & 9 \\ -x_2 - 2x_3 & = & -7. \\ 3x_3 & = & 9 \end{array}$$


These transformations have wrought the system into a state where it can be easily solved.

Indeed, we immediately see that the third equation implies $x_3 = 3$. Substituting into the second equation gives

$$-x_2 - 2 \cdot 3 = -7$$

whose solution is $x_2 = 1$. Finally, knowing the value of x_2 , we can solve the first equation by another substitution. We get

$$x_1 + 6 \cdot 1 = 9,$$

thus $x_1 = 3$ and the triple $(3, 1, 3)$ is the *unique* solution of the system. 

Observant readers might have already identified the ‘kinds’ of transformations that were used in solving the [linear system above](#). Nonetheless, we’re about to spell them out.

The transformations that do not change the solution of a [linear system](#) include

- (1) swapping two equations;
- (2) scaling an equation by a non-zero constant;
- (3) adding a multiple of an equation to *another* equation.

Note that transformations (2) and (3) come with sensible restrictions. Scaling an equation by 0 clearly changes the set of solutions of the system as it basically removes the equation entirely. Adding a multiple of an equation to *itself* suffers from the same problem; it might result in ‘invalidating’ the equation should the scaling factor be -1 .

We now proceed to prove that transformations (1) - (3) truly do not alter the solutions of the initial system.

Theorem 1.1.2 (Gauss-Jordan)

The transformations (1) - (3) of a linear system outlined above do not change its solution set.

PROOF. We will cover transformation (3) here. The proofs for transformations (1) and (2) are similar and thus left as an exercise.

Consider the linear system

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= c_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= c_2 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= c_m \end{aligned}$$

of m equations in variables x_1, \dots, x_n and let (b_1, \dots, b_n) be one of its solutions. Choose a constant k and add the k -multiple of the i -th equation to the j -th equations for some $i, j \in \{1, \dots, m\}$. Hence, the j -th equation of the system gets replaced by

$$(a_{j,1} + k \cdot a_{i,1})x_1 + (a_{j,2} + k \cdot a_{i,2})x_2 + \cdots + (a_{j,n} + k \cdot a_{i,n})x_n = c_j + k \cdot c_i,$$

which can be rearranged to


$$a_{j,1}x_1 + a_{j,2}x_2 + \cdots + a_{j,n}x_n + k \cdot (a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n) = c_j + k \cdot c_i. \quad (1.2)$$

Since (b_1, \dots, b_n) is a solution of the original system, we know that

$$\begin{aligned} a_{i,1}b_1 + a_{i,2}b_2 + \cdots + a_{i,n}b_n &= c_i \\ a_{j,1}b_1 + a_{j,2}b_2 + \cdots + a_{j,n}b_n &= c_j. \end{aligned}$$

Substituting this into equation (1.2) gives

$$c_j + k \cdot c_i = c_j + k \cdot c_i,$$

hence (b_1, \dots, b_n) is also the solution of the transformed system, as required. 

Exercise 1.1.3

Show that transformations (1) and (2) also don't change the set of solutions of the transformed linear system.

Definition 1.1.4 (Elementary operations)

The transformations (1) - (3) outlined above are called *elementary operations* or *row operations*.

As we've seen in [problem 1.1.1](#), the application of transformations (1) - (3) has its purpose in preparing the system for a final back-substitution, where the values of all variables in a row save the first one are known beforehand. A system which is 'ready' to be solved by back-substitution is said to be in *echelon form*.

Definition 1.1.5 (Echelon form)

In each row of a [linear system](#), the first variable with a non-zero coefficient is called the row's *leading variable*.

A linear system is in *echelon form* (or *upper triangular form*) if the leading variable in each row is at least one column to the right of the leading variable in the row above and all rows filled with zeroes are at the bottom.

Example 1.1.6

The system

$$\begin{array}{rcl} x_1 + 6x_2 & = & 9 \\ -x_2 - 2x_3 & = & -7 \\ 3x_3 & = & 9 \end{array}$$

is in echelon form whereas

$$\begin{array}{rcl} 2x_1 + 3x_2 - x_3 & = & 9 \\ 3x_2 - 2x_3 & = & 2 \\ x_1 & - & x_3 = 0 \end{array}$$

is **not**.

For now, we shall employ intuition and a nibble of foresight to guide our transformation of a [linear system](#) into its [echelon form](#). Later, we intend to present a precise algorithm (that computers also use) that achieves this.

Example 1.1.7

We're going to put the system

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

into echelon form and solve it using back-substitution. We'll label the rows of the system by Roman letters and denote transformations accordingly. For example, adding a 3-multiple of row one to row three would be written symbolically as $3 \cdot \text{I} + \text{III}$.

First, we need to get rid of the variable x_1 in rows II and III. This can be done by subtracting

adequate multiples of row I.

$$\begin{array}{rcl} x_1 + x_2 & = & 0 \\ 2x_1 - x_2 + 3x_3 & = & 3 \\ x_1 - 2x_2 - x_3 & = & 3 \end{array} \xrightarrow{\begin{array}{l} -2\text{I} + \text{II} \\ -\text{I} + \text{III} \end{array}} \begin{array}{rcl} x_1 + x_2 & = & 0 \\ -3x_2 + 3x_3 & = & 3 \\ -3x_2 - x_3 & = & 3 \end{array}$$

We continue by subtracting row II from row III.

$$\begin{array}{rcl} x_1 + x_2 & = & 0 \\ -3x_2 + 3x_3 & = & 3 \\ -3x_2 - x_3 & = & 3 \end{array} \xrightarrow{-\text{II} + \text{III}} \begin{array}{rcl} x_1 + x_2 & = & 0 \\ -3x_2 + 3x_3 & = & 3 \\ -4x_3 & = & 0 \end{array}$$

The system is now in **echelon form**. The equation in row III forces $x_3 = 0$. Substitution into row II immediately gives $x_2 = -1$ and one final substitution into row I yields $x_1 = 1$.

Hence, the solution of the system is the triple $(1, -1, 0)$.

Exercise 1.1.8

Using *Gauss-Jordan elimination* solve the systems from examples 1.0.8 and 1.0.9.

All the systems we've studied so far have had the same number of equations as variables. This of course need not be the case in general. Thankfully, Gauss-Jordan elimination can *always* be used to determine the solution set of a **linear system**. However, this set can also be empty or infinite in cases where the number of variables doesn't match the number of equations. The following two examples illustrate this.

Example 1.1.9

The following system has more equations than variables.

$$\begin{array}{rcl} x_1 + 3x_2 & = & 1 \\ 2x_1 + x_2 & = & -3 \\ 2x_1 + 2x_2 & = & -2 \end{array} \tag{1.3}$$

Before we put it into **echelon form** and solve it, let us ponder what the solution set may look like. Intuitively, a linear equation is basically a 'restraint' or 'condition' on the range of possible values the present variables may attain. If there are three equations restraining only two variables, then this restraint may be too harsh and lead to the system having no solution at all. The only case where solution *does* exist involves one of the equations being *redundant* – providing no additional condition. Algebraically, this happens if said equation is a **linear combination** of the other two.

To draw a 'real-life' simile, imagine the price of an apple being \$5/kg and that of bananas \$1.5/kg. Saying that 3 kg of apples and 4 kg of bananas cost, say, \$30 is simply false because we can calculate that this amount actually costs \$21. The third condition on the price of apples and bananas contradicted the previous two; just as a third equation in a **linear system** in two

variables can contradict the first two equations. We tend to call such systems *overdetermined* and will in time dedicate a section to finding a ‘good’ approximation of their solution.

To solve the system (1.3), we transform it into echelon form. First, we subtract twice the first row from the other two.

$$\begin{array}{rcl} x_1 + 3x_2 = 1 & & x_1 + 3x_2 = 1 \\ 2x_1 + x_2 = -3 & \xrightarrow[-2I + III]{-2I + II} & -5x_2 = -5 \\ 2x_1 + 2x_2 = -2 & & -4x_2 = -4 \end{array}$$

Finally, we add $(-4/5)$ -times row II to row III.

$$\begin{array}{rcl} x_1 + 3x_2 = 1 & & x_1 + 3x_2 = 1 \\ -5x_2 = -5 & \xrightarrow{-(4/5)II + III} & -5x_2 = -5 \\ -4x_2 = -4 & & 0 = 0 \end{array}$$

Clearly, the third equation is *redundant* because it provides no condition on the values of the variables. Back-substitution yields $x_2 = 1$ and $x_1 = -2$. As we’ve claimed (but not yet proven), row III is indeed a linear combination of rows I and II. In this particular case, it holds that $(2/5)I + (4/5)II = III$.