

Detailed Syllabus
of the course
Linear Algebra I
for computer science students

JIŘÍ MATOUŠEK

In collaboration with:

Jiří Rohn, Jiří Tůma, Jiří Fiala, Ondřej Pangrác,
Jiří Sgall, Petr Kolman and Milan Hladík

January 13, 2015

Preface

Linear algebra is one of the core subjects for any serious study of mathematics, computer science, physics, and engineering.

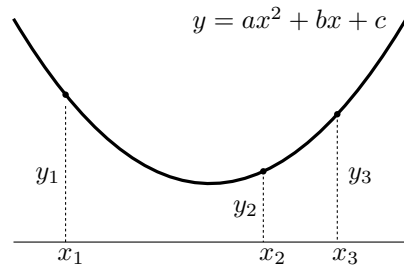
Besides to the factual knowledge, you should master logical reasoning and learn how to express yourselves in mathematics. Linear Algebra is probably the first theory built from the axioms you encounter. Its primary object of study, so called vector space, is defined by several properties (axioms) and everything else is derived from these. Somewhat similar to the rules of chess – giving no description of the shape of a knight, but defining its moves – the definition of a vector space does not describe how a vector looks. It only states the rules for calculations with vectors. Then we can apply the established theory to a wide variety of concrete objects, apparently very different.

Other branches of mathematics are built in this way as well, but linear algebra is quite easy and it is perfectly suitable for demonstration of the development of a mathematical theory. Nevertheless, in time you will find this theory very powerful as well: after mastering basics of linear algebra, you will have no problem to answer questions concerning linear equations, very hard and confounding at first sight and hardly solvable even for mathematically gifted, but unprepared people.

This text is too brief for a proper study of linear algebra and it does not contain proofs. As such, it is not sufficient to **prepare for the exam!** It can be useful when you want to recapitulate the subject and check that you haven't skipped anything important.

1 Systems of Linear Equations

1. Example: To find a quadratic function (in form $y = ax^2 + bx + c$) with its graph passing through given three points gives a system of three linear equations in three variables.

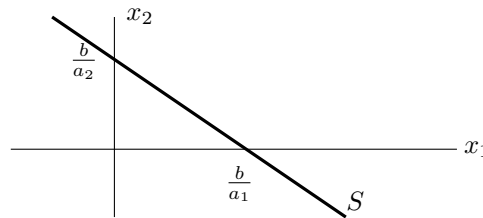


2. The equation $a_1x_1 + a_2x_2 = b$ (1 equation, 2 variables): its solution set is

$$S = \{(x_1, x_2) \in \mathbb{R}^2 : a_1x_1 + a_2x_2 = b\}.$$

Here, \mathbb{R}^2 is the set of all ordered pairs (x, y) , where x, y are real numbers. We will call the ordered pairs, triples, n -tuples of real numbers **vectors**.

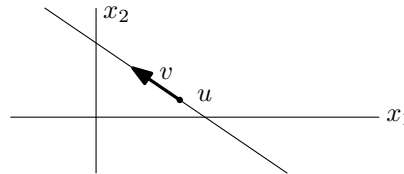
3. Geometrically, the solution set represents a line in the plane (if a_1 and a_2 are not both equal to 0!):



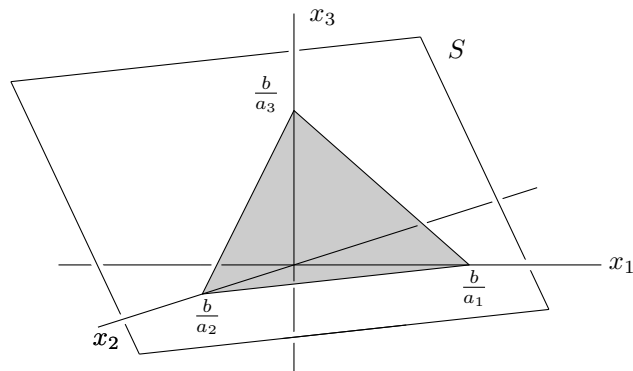
Another way of expressing the same set (parametric form):

$$S = \{\mathbf{u} + t\mathbf{v} : t \in \mathbb{R}\},$$

where \mathbf{u} and \mathbf{v} are suitable vectors from \mathbb{R}^2 .



4. Similarly: the solution set of one linear equation in three variables in form $a_1x_1 + a_2x_2 + a_3x_3 = b$ corresponds to a plane in \mathbb{R}^3 (if a_1, a_2, a_3 are not all equal to 0).



We can express it parametrically as well

$$\{\mathbf{u} + s\mathbf{v} + t\mathbf{w} : s, t \in \mathbb{R}\}$$

for suitable vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ (as we will show later). When we are solving a system of k such equations, we are searching for an intersection of k planes in \mathbb{R}^3 .

5. Generally, we consider a system of m linear equations in n variables in form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

(the first index always denotes the *row*!!). A more concise notation of the same system:

$$A\mathbf{x} = \mathbf{b},$$

where

- A is a **matrix of the system** (matrix with m rows and n columns, or an $m \times n$ (m -by- n) matrix, with the element a_{ij} in the i -th row and the j -th column),

- \mathbf{b} is a column vector of the right-hand side, i.e., an $m \times 1$ -matrix,
- \mathbf{x} is a column vector of the variables, i.e., an $n \times 1$ -matrix.

The notation $A\mathbf{x}$ on the left-hand side is *matrix multiplication*. We will define the multiplication of matrices in general later on.

2 Solving a Linear System: Gaussian Elimination

6. Elementary row operations on a matrix:

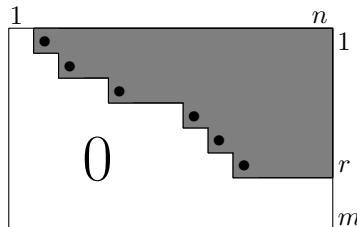
- multiplication of the i -th row by a non-zero number t ,
- addition of the j -th row to the i -th row, $i \neq j$.

By combining the operations (a) and (b) we can get the following

- addition of a t -multiple of the j -th row to the i -th row, $i \neq j$, and
- switching of two rows.

7. **Augmented matrix** of the system $A\mathbf{x} = \mathbf{b}$ is the matrix $(A|\mathbf{b})$, i.e., a matrix A with the column \mathbf{b} added to the right-hand side of the matrix. Claim: elementary row operations of the augmented matrix do not change the solution set of the linear system.

8. **Echelon form of the matrix A** : there exists an index r , $0 \leq r \leq m$, such that the rows $1, 2, \dots, r$ are non-zero, and the rows $r + 1, \dots, m$ equal to zero, and if $j(i) = \min\{j : a_{ij} \neq 0\}$, then $j(1) < j(2) < \dots < j(r)$. (More precisely, this is the definition of the *row echelon form of the matrix*, in the literature sometimes referred to as ‘REF’, as we can analogously define a *column echelon form* analogously. We are not going to talk about it, so we can stick to the shorter term.)



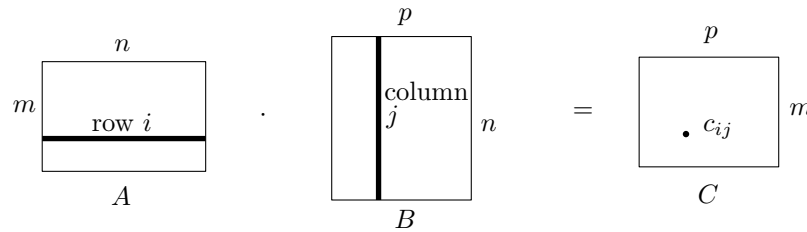
The dots in the picture correspond to the non-zero elements with coordinates $(i, j(i))$, $i = 1, 2, \dots, r$; these are sometimes called **pivots**.

9. **Gaussian elimination:** an algorithm to convert a given matrix A into the echelon form by performing elementary row operations.
10. Solving a system $A\mathbf{x} = \mathbf{b}$ by elimination: the matrix A is converted to the echelon form while all the row operations are applied to the augmented matrix. How does the solution set of a system look like, when the matrix A is in the echelon form? If b_{r+1}, \dots, b_m are not all equal to zero, then there is no solution. Otherwise we get all the solutions when we choose variables x_j in the columns not containing the pivot arbitrarily (there are $n - r$ such variables) and calculate (unambiguously) values of the r remaining variables. In the special case when $r = n$ there is exactly one solution.
11. Numerical issues, ill-conditioned matrices (a small change of the matrix causes huge change in the solution). Example (2×2): geometric interpretation (almost parallel lines).

3 Operations on Matrices; Special Types of Matrices

12. Addition of matrices (of the same type!) element-wise, multiplication by a real number element-wise.
13. **Transpose of a matrix** A^T : the element a_{ij} goes to position (j, i) .
Symmetric matrix: a square (i.e., $n \times n$) matrix, $A^T = A$.
14. **Identity matrix** I_n (n -by- n , ones in positions (i, i) , $i = 1, 2, \dots, n$, zeros elsewhere).
15. Matrix A is **diagonal** if all non-zero elements are on the main diagonal, i.e., $a_{ij} = 0$ for all $i \neq j$.
16. **Matrix multiplication:** the product AB is not defined for every pair of matrices A and B , but only if the number of columns of A equals to the number of rows of B , i.e., A is an $m \times n$ -matrix and B is an $n \times p$ -matrix. The product AB is then an $m \times p$ -matrix C , where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}.$$



Check: $AI_n = I_m A = A$, for every $m \times n$ -matrix A .

17. Multiplication and transposition: $(AB)^T = B^T A^T$ (more precisely: the product AB is defined if and only if the product $B^T A^T$ is defined, and in that case the equality holds – similar assumptions apply to further matrix equalities below as well.)
18. Distributivity: $A(B + C) = AB + AC$, and similarly from the right.
19. Matrix multiplication is associative.
20. Let A be an $n \times n$ -matrix. The matrix B is an **inverse** of A if $AB = I_n$. (Be careful, only a square matrix can have an inverse!) The inverse of a matrix A , if it exists, is referred to by A^{-1} .
21. Which matrix has an inverse? We need the following notion for an answer: A square matrix A is **non-singular** if the system $A\mathbf{x} = \mathbf{0}$ has a unique solution (i.e., $\mathbf{x} = \mathbf{0}$).
22. Theorem: An $n \times n$ -matrix A has an inverse if and only if it is non-singular. In that case the inverse is determined uniquely and the following holds: $AA^{-1} = A^{-1}A = I_n$, i.e., the inverse is a left-inverse and a right-inverse as well. Non-singular matrices are also called **invertible**.
23. In the proof and not only there we can use the following claim: A matrix is non-singular \Leftrightarrow in (some) echelon form there is $r = n \Leftrightarrow$ the system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every \mathbf{b} .
24. Multiplication and inversion: $(AB)^{-1} = B^{-1}A^{-1}$ (as with transposition).
25. Calculating an inverse of a matrix: We take the matrix $(A|I_n)$ and we convert it by row operations to the form $(I_n|B)$ (if possible) – then $B = A^{-1}$. If it is impossible, A is singular.

26. Elementary row operations on a matrix correspond to multiplying the matrix by a suitable square non-singular matrix from the left. A product of invertible matrices is invertible and therefore a sequence of elementary row operations corresponds to multiplication of the matrix from the left by a suitable non-singular matrix.

4 Groups and Permutations

27. Now we step aside from the main topic of linear algebra and explore two important mathematical structures – groups and fields. This will be our first encounter with the *abstract approach* in mathematics, where objects are defined by axioms (“rules of the game”).

28. If X is a set, a **binary operation** on X is an arbitrary mapping $X \times X \rightarrow X$.

Informally, a binary operation assigns an element of X to every pair of elements $a, b \in X$, as a result of the operation applied on a and b .

29. We can view binary operations as a generalization of the “four basic arithmetic operations” – addition, subtraction, multiplication and division. Addition, subtraction, and multiplication are indeed examples of binary operations on the set \mathbb{R} of all real numbers. (But there are many more interesting examples of binary operations and they do not have to concern numbers at all.)
30. *Warning:* the division is not a binary operation on the set \mathbb{R} (but it is a binary operation on the set $\mathbb{R} \setminus \{0\}$). The subtraction is not a binary operation on the set of all natural numbers.
31. Binary operations are usually denoted by symbols $\circ, *, +$, etc. The notation is similar to the basic arithmetic operations, i.e., $a \circ b$ denotes the result of a binary operation \circ applied on a and b .
32. Here are two important properties that a binary operation may or may not have:

A binary operation \circ on a set X is called **commutative** if $a \circ b = b \circ a$ for all $a, b \in X$, and it is called **associative** if $a \circ (b \circ c) = (a \circ b) \circ c$ for all $a, b, c \in X$.

33. Examples: Addition $+$ on \mathbb{R} is associative, as well as commutative. Subtraction $-$ on \mathbb{R} is neither associative, nor commutative (check!).
34. One of the most important objects in the whole mathematics is a **group**. It is defined by **axioms**, i.e., properties required from it.

A **group** is a pair (G, \circ) where G is a set and \circ is a binary operation on G , satisfying the following axioms:

- (A) The operation \circ is associative.
- (E) There exists an element $e \in G$ such that $a \circ e = e \circ a = a$ for every $a \in G$. (Such e is called the **identity element** of the group, sometimes also the **neutral element**.)
- (I) For every $a \in G$ there is a $b \in G$ such that $a \circ b = b \circ a = e$, where e is the identity element. (Such b is denoted by a^{-1} and is called the **inverse element** of the element a .)

35. Notes:

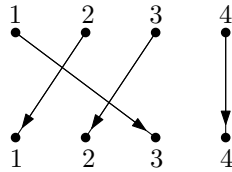
- *Beware*, the definition of a group also implies the requirement that for every $a, b \in G$, also $a \circ b \in G$ (given by the definition of a binary operation).
- Instead of “group (G, \circ) ” it is usually enough to say only “group G ”, if it is clear, which operation is used.
- The notation $a \circ b$ is often shortened to ab (like for multiplication), again only if the operation is obvious from the context.

36. We can derive many other properties of groups from the axioms. Examples of corollaries: There is only one identity element in every group. There is exactly one inverse element for every element a in any group. In a group, we can use “cancellation”, i.e., $a \circ c = b \circ c$ yields $a = b$.
37. We need to derive *every* property of a group from the axioms. The fact that something is true for one concrete group, or even for many different groups, by no means implies that the thing is true in every group.
38. Even though the group axioms look simple, the universe of groups is very complicated and even after a hundred years of study, it conceals

many secrets. Only by the end of the twentieth century the so called “enormous theorem” was proved. (Roughly speaking, it describes all possible finite “building blocks” of groups.) Its proof consists of several *thousands* pages and one concrete example of a group described by the enormous theorem, the so called *monster*, has approximately 8×10^{53} elements. (Do not worry, linear algebra is easier than the group theory and we will explain only very simple things concerning groups.)

39. What are groups good for? In mathematics, they show up in the proof of impossibility of a general solution of a quintic equation by algebraic operations, in number theory, in enumerative combinatorics, and many other fields. In physics, the symmetry conditions of physical laws are usually essential, and these symmetries are described by suitable groups. Groups are used in crystallography, cryptography, image analysis and other fields. We will see some use of groups in linear algebra as well.
40. One more notion: A **subgroup** of a group (G, \circ) is a subset $H \subseteq G$ such that $e \in H$ (where e is the identity element of G), $a^{-1} \in H$ for every $a \in H$, and $a \circ b \in H$ for every $a, b \in H$. That is, H forms a group under the operation “inherited” from G .
41. Examples of groups (and subgroups):
- $(\mathbb{R}, +)$; $(\mathbb{Z}, +)$ (where \mathbb{Z} is the set of all integers); the set of all positive rational numbers with multiplication; the set $\{-1, 1\}$ with multiplication. In all these cases the operation is commutative and we are talking about **commutative**, or **Abelian**, groups.
 - $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{R}, +)$; $(\mathbb{N}, +)$ *is not* a subgroup of $(\mathbb{Z}, +)$ (since it is not a group).
 - For every n , the set of all *invertible* $n \times n$ -matrices together with the operation of multiplication forms a group. For $n \geq 2$, this group *is not* commutative. (On the other hand, the set of all $n \times n$ -matrices with multiplication is not a group.)
 - The set of all rotations about origin in three-dimensional space together with the operation of composition forms a group, which is non-abelian as well (if you rotate, say, a cup 90° around the x -axis and then 90° around the z -axis, it will be in other position than if you make a 90° -turn around the z -axis followed up by 90° -turn around the x -axis – try it with an empty cup).

42. Permutations are another rich source of examples. Let us recall: A **permutation** of a set X is a one-to-one correspondence (bijection) $X \rightarrow X$. Let S_n be the set of all permutations of the set $\{1, 2, \dots, n\}$.
43. We use the two-line notation for permutations, e.g., $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$, or a figure with arrows (a bipartite graph):




44. Permutations can be composed as maps; for $p, q \in S_n$ the composition $p \circ q$ is defined as $p \circ q(i) = p(q(i))$, $i = 1, 2, \dots, n$. The set S_n together with the operation \circ forms a group, called the **symmetric group**.
45. For $n \geq 3$, the group S_n is non-abelian.
46. Subgroups of the symmetric group are called **permutation groups**.
47. In the course of time we will need the *sign of a permutation*. First we define the set of **inversions** of a permutation p :

$$I(p) = \{(i, j) : i < j \text{ and } p(i) > p(j)\}.$$

Interpretation: a crossing of arrows in the two-line notation of p . The **sign of permutation** is then $\text{sgn}(p) = (-1)^{|I(p)|}$.

48. Claim (composition of permutations and sign): $\text{sgn}(p \circ q) = \text{sgn}(p) \text{sgn}(q)$.
Proof: a figure with arrows.
49. A **transposition** is a permutation switching two elements and leaving all others in place. Claim: The sign of any transposition is -1 . Every permutation is a composition of transpositions.
50. In the language of group theory, we can formulate the last claim as follows: The set $T \subseteq S_n$ of all transpositions generates the group S_n . A general definition: Let G be a group and $M \subseteq G$ an arbitrary subset of G . We say that M **generates** the group G (or that M is a *set of generators* of G) if the only subgroup of G containing M is the whole of G .

51. Equivalently, every $a \in G$ can be expressed by finitely many elements of M using the group operation and the inversion. (Delicacy: $M = \emptyset$ generates the group $\{e\}$ formed only by the identity element.)
52. Some popular puzzles are in fact permutation groups in disguise. The *15-puzzle* is a frame of 4×4 cells with tiles numbered 1 through 15 and a free space allowing to move the tiles (horizontally or vertically).



1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

Around the year 1880, hundreds of thousands people were trying to solve the puzzle to move the tiles positioned as in the figure to the same position with only 14 and 15 switched. It has no solution, as one can show using the sign of a permutation.

53. A more modern group puzzle is the well-known *Rubik's cube*. With that we want to express a given element of a certain permutation group (cube with mixed colours on the faces) by generators (rotation of the faces). (Recently it has been proved by extensive calculations that every position can be solved in at most 20 moves and for some positions 20 moves are needed.)

5 Fields (in Algebra)

54. We can apply “elementary arithmetic operations” on the rational, real, and complex numbers; operation of addition and multiplication, and derived (inverse) operations of subtraction and division.
55. A field is an algebraic structure with defined operations with similar properties (and thus we can “calculate” with its elements in a way similar to the real numbers). It is defined by axioms again.

A **field** is a set \mathbb{K} together with two binary operations $+$ (addition) and \cdot (multiplication), satisfying the following axioms:

- (SG) The set \mathbb{K} with the operation $+$ forms an *Abelian group*. The identity element of this group is denoted by 0 and the inverse element to a is denoted by $-a$.
- (NG) The operation \cdot is commutative, and the set $\mathbb{K} \setminus \{0\}$ with this operation (strictly speaking with its restriction to $\mathbb{K} \setminus \{0\}$) forms a *group*. The identity element of this group is called 1 and the inverse element to a is denoted by a^{-1} .
- (D) Multiplication is **distributive** over addition, i.e. $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ for every $a, b, c \in \mathbb{K}$.

For the multiplication $a \cdot b$, we usually use a shorter notation ab . The subtraction is defined as $a - b = a + (-b)$, and division as $a/b = a \cdot b^{-1}$.

Historically, fields were called *commutative fields*, as the expression “field” was used for *division rings*. That is a structure satisfying all axioms of a field, only commutativity of multiplication is not assumed. We will always understand the field as commutative.

- 56. Claims on matrix multiplication, inverses, or solutions of systems of linear equations hold for any field. We do not need to work with numbers. Everything needs to be proved from the axioms (using nothing else!!!).
- 57. Examples of fields: rational numbers \mathbb{Q} , real numbers \mathbb{R} , complex numbers \mathbb{C} , two-element \mathbb{Z}_2 . More exotic: $\mathbb{R}(x)$ – elements are all rational functions $p(x)/q(x)$, where $p(x)$ and $q(x)$ are polynomials with real coefficients.
- 58. Notation \mathbb{Z}_n (residue classes modulo n , represented by numbers $0, 1, \dots, n-1$, with operations of addition and multiplication modulo n). \mathbb{Z}_3 is a field, \mathbb{Z}_4 IS NOT!!!
- 59. Claim: \mathbb{Z}_n is a field if and only if n is a prime. Idea of the proof: If n is a composite number in the form $n = k\ell$, then residue classes of k and ℓ are *zero divisors*, i.e., their product is 0 in \mathbb{Z}_n . If n is a prime, we need to show that for every non-zero $\ell \in \mathbb{Z}_n$: the map ‘multiplication

by $\ell': \mathbb{Z}_n \rightarrow \mathbb{Z}_n$ is surjective (onto). Trick: check that the function is injective (one-to-one).

60. Notation: $\text{GF}(q)$, a finite field with q elements (Galois Field) exists if and only if q is a power of a prime, and then there is exactly one (without proof). Finite fields are very important in computer science (e.g., for codes, on computer discs or DVDs).

61. The **characteristics** of a field: the smallest $n \geq 1$ such that

$$\underbrace{1 + 1 + \cdots + 1}_{n\text{-times}} = 0,$$

or 0 if there is no such n . Claim: characteristic is always a prime or 0.

6 Vector Spaces

62. So far, we took vectors as ordered n -tuples of real numbers in the form $\mathbf{v} = (v_1, \dots, v_n)$, living in \mathbb{R}^n (Cartesian product of n copies of \mathbb{R} ; e.g., \mathbb{R}^2 describes the plane). We can add them together and multiply them by a real number. In the same way as we have generalized the real numbers to fields by axioms, we generalize \mathbb{R}^n to so called vector spaces.

63. We can say that linear algebra is a study of vector spaces. When we talk about vector spaces, we can always visualize \mathbb{R}^2 , \mathbb{R}^3 and \mathbb{R}^n in general as the basic (and most important) examples.

A **vector space** over a field \mathbb{K} is a set V (elements = **vectors**) together with a binary operation $+$ (vector addition) and an operation \cdot (multiplication of a vector by scalar from the field \mathbb{K} ; it is a map $\mathbb{K} \times V \rightarrow V$) satisfying the following axioms:

- (SG) The set V together with the operation $+$ is an *Abelian group*. Its neutral element is called $\mathbf{0}$, and the inverse of the vector \mathbf{v} is called $-\mathbf{v}$. [Be aware that we have two distinct zeros, 0 in \mathbb{K} and $\mathbf{0}$ in V !!!]
- (NA) Scalar multiplication of vectors is ‘associative’, i.e., $a \cdot (b \cdot \mathbf{v}) = (a \cdot b) \cdot \mathbf{v}$ for every $a, b \in \mathbb{K}$ and every $\mathbf{v} \in V$.
- (N1) We have $1 \cdot \mathbf{v} = \mathbf{v}$ for every $\mathbf{v} \in V$ (and $1 \in \mathbb{K}$ is the identity of the field).
- (D1) The following distributivity holds: $(a+b) \cdot \mathbf{v} = (a \cdot \mathbf{v}) + (b \cdot \mathbf{v})$, for every $a, b \in \mathbb{K}$ and every $\mathbf{v} \in V$,
- (D2) and also this distributivity: $a \cdot (\mathbf{u} + \mathbf{v}) = (a \cdot \mathbf{u}) + (a \cdot \mathbf{v})$, for every $a \in \mathbb{K}$ and every $\mathbf{u}, \mathbf{v} \in V$.

Instead of $a \cdot \mathbf{v}$, we write the shorter $a\mathbf{v}$. Note that for any $\mathbf{u}, \mathbf{v} \in V$ and $a \in \mathbb{K}$ we also have $\mathbf{u} + \mathbf{v} \in V$ and $a\mathbf{v} \in V$.

64. Examples:

- $\{\mathbf{0}\}$ (trivial vector space).
- \mathbb{K}^n (**arithmetic vector space** of dimension n over \mathbb{K}) for any field \mathbb{K} .
- The set of all m -by- n matrices with elements from \mathbb{K} (or any other fixed m -by- n).
- $\mathbb{R}[x]$ (all polynomials with real coefficients).
- Polynomials of degree at most 293 with real coefficients (or any other given maximum degree).
- The set of all subsets of a set X as a vector space over $\text{GF}(2)$ (addition = symmetric difference of the sets).
- The set of all functions $\mathbb{R} \rightarrow \mathbb{R}$ ($(f+g)(x) = f(x) + g(x)$ etc.), similarly the set of all *continuous* functions $\mathbb{R} \rightarrow \mathbb{R}$ or of all *differentiable* functions $\mathbb{R} \rightarrow \mathbb{R}$.

- An exotic example: \mathbb{R} (real numbers) as a vector space over \mathbb{Q} (rat. numbers).
65. Claims on vector spaces: $0\mathbf{x} = \mathbf{0}$, $(-1)\mathbf{x} = -\mathbf{x}$, $a\mathbf{x} = \mathbf{0}$ if and only if $a = 0$ or $\mathbf{x} = \mathbf{0}$.
66. A **subspace** of a vector space V is a subset $W \subseteq V$, which is a vector space in respect to $\mathbf{0}$, “+” and “.” inherited from V . That is, $\mathbf{0} \in W$, $\mathbf{u} + \mathbf{v} \in W$ for every $\mathbf{u}, \mathbf{v} \in W$, and also $a\mathbf{v} \in W$ for every $a \in \mathbb{K}$ and every $\mathbf{v} \in W$.
67. Example: vector subspaces \mathbb{R}^2 are (geometrically) origin, the whole \mathbb{R}^2 , and every line passing through the origin (we will check later).
68. Observation: the intersection of an arbitrary system of subspaces of a vector space V is again a subspace. Definition: If X is a subset of a vector space V , the **subspace generated by X** is the intersection of all subspaces W of V containing X . Notation: $\text{span}(X)$ (in the literature also $\langle X \rangle$, $\mathcal{L}(X)$, $[X]$, called also **linear span** or **linear hull of X**).
69. If $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ are vectors, every expression $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_n\mathbf{v}_n$, where $a_i \in \mathbb{K}$, is called **linear combination $\mathbf{v}_1, \dots, \mathbf{v}_n$** (in a linear combination, we always have a *finite* number of vectors!). The vector $\mathbf{0}$ is considered to be a linear combination of an empty set of vectors. Claim (explicit description of a subspace generated by X): $\text{span}(X)$ is a set of all linear combinations of vectors from X .
70. Let A be an m -by- n matrix. Vector spaces related to it:
- **row space** (= a subspace of \mathbb{K}^n generated by the rows of A),
 - **column space** (= a subspace of \mathbb{K}^m generated by the columns of A),
 - **kernel** or **null space** (= a subspace of \mathbb{K}^n generated by all solutions of the system $A\mathbf{x} = \mathbf{0}$), notation: $\text{Ker } A$.

Observation: elementary row operations on a matrix do not change its row space or its kernel.

7 Linear Dependence, Basis, Dimension

71. A collection (finite sequence) of vectors $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ is **linearly independent** if from the equality $a_1\mathbf{v}_1 + \dots + a_n\mathbf{v}_n = \mathbf{0}$ follows that $a_1 = a_2 = \dots = a_n = 0$, i.e. vectors can be combined to equal zero in only one possible, trivial way.

(The vectors in a collection, in contrast to a set, can repeat, but as soon as $\mathbf{v}_i = \mathbf{v}_j$, the collection is linearly dependent.)

72. Infinite collection of vectors is linearly independent, if every finite sub-collection is linearly independent. (What is an infinite collection? Similar to a set, but the elements can repeat themselves; formally we write an infinite collection as $(\mathbf{v}_i)_{i \in I}$, where I is an infinite set of ‘indices’.)
73. Examples of linearly independent collections:
- $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ – rows of the identity matrix I_n (i.e., so called **standard basis** of \mathbb{R}^n);
 - first r rows of a matrix in the echelon form;
 - $(x^i)_{i=0,1,\dots}$ in $\mathbb{R}[x]$,
 - $(1, \sqrt{2})$ in \mathbb{R} as vector space over \mathbb{Q} .

74. Alternative, but maybe more intuitive description of linear independence: $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ is linearly independent if every \mathbf{v}_i “adds something” to the linear span: $\mathbf{v}_i \notin \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n)$ for every $i = 1, 2, \dots, n$.

75. Definition: Let B be a collection of vectors in a vector space V ; it is called a **generating system** of V if $\text{span}(B) = V$.

A linearly independent generating system of a vector space V is called a **basis** of the space V .

76. Examples: an empty system is a basis of the trivial space $\{\mathbf{0}\}$; $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ is a basis of \mathbb{K}^n ; $(1, x, x^2, \dots)$ is a basis of $\mathbb{R}[x]$.
77. Claim: A minimal generating system (i.e., no proper subsystem generates the whole space) is a basis. Therefore we can select a basis from any finite generating system.

78. Theorem: *every vector space has a basis*. The proof requires the axiom of choice. We have shown (only) for spaces with some finite generating system (such spaces are called **finitely generated**).
79. Can one vector space have two bases of different sizes? NO!! For the proof, we need the **Steinitz exchange lemma**.
80. First, an **exchange lemma**: If $G = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ is a generating system of a space V , $\mathbf{w} \in V$ is a vector, and $\mathbf{w} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_n\mathbf{v}_n$ is its expression in terms of vectors from G , then whenever $a_i \neq 0$, the system $(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{w}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n)$ is a generating system as well (i.e., the vector \mathbf{v}_i with a non-zero coefficient can be replaced by \mathbf{w}).
81. **Steinitz exchange lemma**: If $N = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m)$ is a linearly independent collection of vectors in V and $G = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ is a generating system of V , then $m \leq n$ and we can replace some m vectors from G by the vectors $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ so that we get a generating system again.
82. Main corollary: All bases of a finitely generated space are finite and they have the same number of vectors. (In an arbitrary vector space, every basis has the same cardinality, but we will not prove this.)

The **dimension** of a vector space V is the cardinality of a (and therefore any) basis of V .

83. Another corollary of the Steinitz lemma: An arbitrary linearly independent system in a finitely generated space can be extended to a basis.
84. Then: If W is a subspace of a finitely generated space V , then
- $$\dim(W) \leq \dim(V)$$
- (and in particular, W is finitely generated). In case of equality we have $W = V$.
85. Example: what are the subspaces of \mathbb{R}^2 ? They can have dimension 0 (then it is $\{\mathbf{0}\}$), 2 (then it is \mathbb{R}^2), or 1, and a one-dimensional vector space consists of all multiples of a non-zero vector, that means it is a line passing through $\mathbf{0}$. Similarly for \mathbb{R}^3 : there will be additional planes passing through $\mathbf{0}$.
86. Terminology: **coordinates of a vector relative to a given basis**.

8 Finding a Basis, Rank of a Matrix

87. How can we calculate the dimension (and find a basis) of a space? Let $V = \text{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ be a vector space spanned by $\mathbf{a}_1, \dots, \mathbf{a}_m$ – given vectors from \mathbb{K}^n . We enter $\mathbf{a}_1, \dots, \mathbf{a}_m$ as rows of a matrix A (then V is its row space). *Gaussian elimination* is an algorithm for finding a basis: non-zero rows of some echelon form make up a basis of V .

The **rank of a matrix** A is defined as the dimension of its row space, and we will denote it by $\text{rank } A$.

The rank also equals to the number of non-zero rows in an echelon form (and therefore this number does not depend on the progress of the Gaussian elimination, which is not obvious from the algorithm itself).

88. Theorem (one of the “marvels” of linear algebra): The rank of a matrix is equal to the dimension of the column space as well.

Proof:

- Obvious for the reduced echelon form.
- Elementary row operations, and more general multiplication by an invertible matrix R from the left, do not change the *dimension* of the column space (even when the column space changes). This follows from the claim: if $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ is a basis of the column space of A , then $\{R\mathbf{v}_1, \dots, R\mathbf{v}_r\}$ spans the column space of RA .

89. From the echelon form, we can also find the basis of $\text{Ker}(A)$, and deduce that

$$\dim(\text{Ker } A) + \text{rank}(A) = n$$

for every matrix A with n columns. This is quite an important equation.

90. $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$ (A, B are matrices for which the product AB is defined). Because: the row space of $AB \subseteq$ the row space of B , and the column space of $AB \subseteq$ the column space of A .
91. From this: $\text{rank}(RA) = \text{rank}(A)$ for a (square) invertible R .

9 Linear Maps

92. A map $f: U \rightarrow V$, where U and V are vector spaces (over the same field!), is **linear** if $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$ and $f(a\mathbf{u}) = af(\mathbf{u})$ for every $\mathbf{u}, \mathbf{v} \in U$ and $a \in \mathbb{K}$.

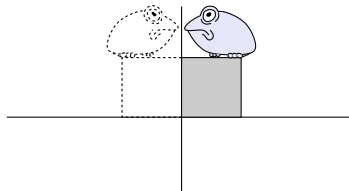
93. A composition of linear maps is a linear map too (if they can be composed!).

94. Example (simple): linear maps $\mathbb{R}^1 \rightarrow \mathbb{R}^1$ must be in the form $x \mapsto ax$, $a \in \mathbb{R}$.

95. Linear maps $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ are already quite interesting. Examples:

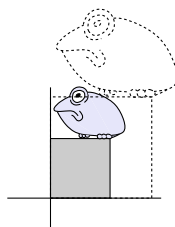
- projection to the x -axis,
- projection to a given line passing through $\mathbf{0}$,
- reflection, e.g.,

$$(x, y) \mapsto (-x, y)$$



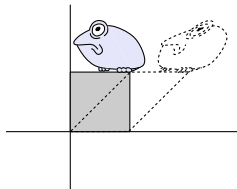
- enlargement (homothety), e.g.,

$$(x, y) \mapsto (1.7x, 1.7y)$$



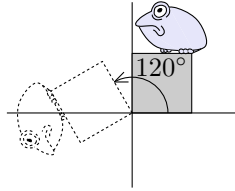
- shearing, e.g.,

$$(x, y) \mapsto (x + y, y)$$

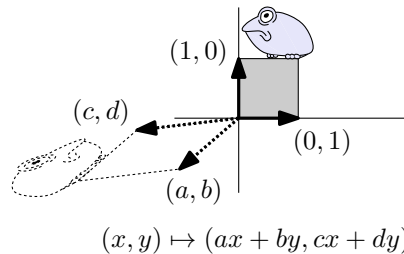


- **rotation** about $\mathbf{0}$, e.g.,

$$(x, y) \mapsto \left(-\frac{1}{2}x - \frac{\sqrt{3}}{2}y, \frac{\sqrt{3}}{2}x - \frac{1}{2}y\right)$$



96. General form: $f(x, y) = (ax + by, cx + dy)$, there are no others. Matrix form: $f(\mathbf{v}) = A\mathbf{v}$, where $\mathbf{v} \in \mathbb{R}^2$ is a column vector (x, y) and A is a matrix with rows $(a, b), (c, d)$.



97. Claim (Every choice of values on the basis determines the linear map uniquely) Let U, V be vector spaces and B a basis of U . For every map $f: B \rightarrow V$, there is exactly one linear map $\bar{f}: U \rightarrow V$ such that $\bar{f}(b) = f(b)$ for every $b \in B$.
98. From that: when we know (geometrically) that, e.g., a rotation about $\mathbf{0}$ by an angle τ is a linear map, we can easily express it; we get $(x, y) \mapsto (x \cos \tau - y \sin \tau, x \sin \tau + y \cos \tau)$.
99. Example: Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be vertices of a regular n -gon with the origin in $\mathbf{0}$. Show that $\mathbf{s} = \sum_{i=1}^n \mathbf{v}_i = \mathbf{0}$. An elegant solution: let τ be a rotation about $\mathbf{0}$ by the angle $\frac{2\pi}{n}$, then $\tau(\mathbf{s}) = \mathbf{s}$, and therefore $\mathbf{s} = \mathbf{0}$.
100. An arbitrary linear map $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is in the form $f(\mathbf{x}) = A\mathbf{x}$, where \mathbf{x} is a column vector from \mathbb{R}^n and A is an m -by- n matrix; its *columns* are the images of the basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$. Matrices of the usual geometrical transformations, like a rotation about origin, appear for example in computer graphics.
101. Terminology: **matrix of a (linear) map $f: U \rightarrow V$ relative to given bases** of spaces U and V ; the j -th column of this matrix is the coordinates of the image of the j -th vector of the basis of the space U relative to the basis of the space V .

102. *Composition of linear maps and matrix multiplication* : If V_1, V_2, V_3 are vector spaces and B_i is a basis of V_i , $f: V_2 \rightarrow V_1$ is a linear map with a matrix A relative to the bases B_2 and B_1 , and $g: V_3 \rightarrow V_2$ is a linear map with a matrix B relative to the bases B_3 and B_2 , then $f \circ g: V_3 \rightarrow V_1$ has the matrix AB relative to the bases B_3 and B_1 . Proof from the associativity of matrix multiplication. Let $\mathbf{v} \in V_3$, and let \mathbf{x} be the coordinate vector of \mathbf{v} with respect to B_2 . Then $g(\mathbf{v})$ has the coordinates $B\mathbf{x}$ and $f(g(\mathbf{v}))$ has the coordinates $A(B\mathbf{x}) = (AB)\mathbf{x}$.
103. Example: matrix multiplication of rotations about origin in \mathbb{R}^2 yields the addition formulas for sine and cosine functions.
104. If B and C are two bases of the space V , then the matrix of the identity map $\text{id}: V \rightarrow V$ relative to bases B and C is called a **change of basis matrix** from B to C . If \mathbf{x} is a coordinate vector of some $\mathbf{v} \in V$ relative to the basis B , then the coordinates of \mathbf{v} in the basis C are given by the vector $A\mathbf{x}$, where A is the change of basis matrix from B to C .

10 Isomorphism of Vector Spaces

105. What does it mean that the vector spaces V and W are “the same”? There exists an **isomorphism** $f: V \rightarrow W$ between them, which is a linear map with an inverse map that is also linear (equivalently if f is an isomorphism if it is linear and bijective). An isomorphism is something like renaming the vectors: vectors in isomorphic spaces can “look” different, but they “behave” exactly in the same way.
106. An isomorphism maps a basis to a basis, and therefore it preserves the dimension.

107. Claim (there is only one n -dimensional vector space over \mathbb{K}): every n -dimensional vector space V over \mathbb{K} is isomorphic to \mathbb{K}^n .

Proof: select a basis of V ; an isomorphism $V \rightarrow \mathbb{K}^n$ maps the vector $\mathbf{v} \in V$ to its coordinates in this basis. (Note: many isomorphisms = many “possible views” of the given vector space!)

108. If $\dim(U) = \dim(V) = n$, $f: U \rightarrow V$ is linear, and A is the matrix of f relative to some bases, then f is an isomorphism if and only if A is invertible. (From this, we get another proof of the theorem on matrix inverses from item 3).

11 Affine Subspaces

109. *Affine subspaces:* A subset F of a vector space V that is either empty or in the form $F = \mathbf{x} + U = \{\mathbf{x} + \mathbf{u} : \mathbf{u} \in U\}$, where U is a (vector) subspace of V , is called an **affine subspace** of V .
110. We have $U = \{\mathbf{u} - \mathbf{v} : \mathbf{u}, \mathbf{v} \in F\}$, and therefore F determines U . The **dimension** of F is defined as $\dim(U)$. For example, general lines and planes in \mathbb{R}^3 are affine subspaces. Terminology: a one-dimensional affine subspace is called a **line**, a two-dimensional a **plane**, and an $(n - 1)$ -dimensional affine subspace of an n -dimensional space is called a **hyperplane**.
111. If $f: U \rightarrow V$ is a linear map and $\mathbf{b} \in V$ is a given vector, then $f^{-1}(\mathbf{b})$ is an affine subspace of U ; if it is not empty, it has the form $\mathbf{x} + \text{Ker}(f)$, where \mathbf{x} is an (arbitrary) vector satisfying $f(\mathbf{x}) = \mathbf{b}$.
112. The same in the matrix language: set of all solutions of the system $A\mathbf{x} = \mathbf{b}$, where A is an m -by- n matrix and \mathbf{b} is an m -component vector, is either empty, or of the form $\mathbf{x}_0 + L$, where \mathbf{x}_0 is an arbitrary solution of the system $A\mathbf{x} = \mathbf{b}$ and L is the set of all solutions of the **homogeneous** system $A\mathbf{x} = \mathbf{0}$. Finding all solutions of the system $A\mathbf{x} = \mathbf{b}$: we find one solution \mathbf{x}_0 (if there is any) and some basis for the space of solutions of the homogeneous system $A\mathbf{x} = \mathbf{0}$, i.e., $\text{Ker}(A)$.
113. Summary of our knowledge about solutions of systems of linear equations $A\mathbf{x} = \mathbf{b}$, and different views of it:
- Vector-space view: is \mathbf{b} in the subspace spanned by the columns of A ?
 - Geometric view: an intersection of hyperplanes in \mathbb{K}^n .
 - Linear-mapping view: the preimage of the vector \mathbf{b} under linear map $\mathbf{x} \mapsto A\mathbf{x}$; the solution is an affine subspace of \mathbb{K}^n .