19 Review of linear algebra

The content of our ODE course requires some basic knowledge from the linear algebra course, and this lecture serves as a quick review of the pertinent material. I also introduce the notation that I use throughout the rest of the lectures.

19.1 Matrix arithmetics

Matrix of the size m by n is a rectangular table with m rows and n columns. I usually denote matrices with bold capital letters $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \ldots$ and the corresponding elements of these matrices with small letters $a_{ij}, b_{ij}, c_{ij}, \ldots$ a_{ij} means that this element is on the intersection of the i-th row and the j-th column:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ \vdots & \ddots & & & \vdots \\ a_{i1} & & a_{ij} & & a_{in} \\ \vdots & & & \ddots & \vdots \\ a_{m1} & \ldots & a_{mj} & \ldots & a_{mn} \end{bmatrix}.$$

Note that I use square brackets to denote matrices, very often curvy brackets are used:

$$\boldsymbol{B} = \begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix}.$$

The meaning of this notation is absolutely the same. However, this should not be confused with

$$\begin{vmatrix} -1 & 0 \\ 0 & 5 \end{vmatrix},$$

which denotes the determinant of a matrix. I often use the shortcut notation $\boldsymbol{A} = [a_{ij}]_{m\times n}$ to specify matrix $\boldsymbol{A}$. Sometimes the dimensions of the matrix are convenient to write as indices: $\boldsymbol{A}_{m\times n}$ denotes a matrix with m rows and n columns.

The matrix *transpose* exchanges rows and columns of a matrix and is usually denoted as $\boldsymbol{A}^\top$. Formally, transpose of a matrix $\boldsymbol{A}_{m\times n}$ is the matrix $\boldsymbol{B}_{n\times m} = \boldsymbol{A}^\top_{m\times n}$ such that $b_{ij} = a_{ji}, i = 1, \ldots, n, j = 1, \ldots, m$. Here is a simple example:

$$\begin{bmatrix} -1 & 1 & 3 \\ 2 & 0 & -2 \end{bmatrix}^\top = \begin{bmatrix} -1 & 2 \\ 1 & 0 \\ 3 & -2 \end{bmatrix}.$$

There are a lot of special matrices, which are useful to remember.

- The matrix with only one row is called the *row-vector*:

$$\boldsymbol{a} = (a_1, \ldots, a_n).$$

MATH266: Intro to ODE by Artem Novozhilov, e-mail: artem.novozhilov@ndsu.edu. Spring 2024

- The matrix with only one column is called the *column-vector*:

$$\boldsymbol{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

  Quite obvious, a transposed row-vector becomes a column-vector and vice versa. To denote vectors I use small bold letters, in class I usually write $\vec{a}, \vec{b}$. Frequently I do not specify whether it is a row or a column vector, in this case it is usually meant to be a column-vector. For vectors I use only one index. Depending on the nature of constants in vectors (they can be either real or complex), I use the notations $\boldsymbol{a} \in \mathbf{R}^n, \boldsymbol{b} \in \mathbf{C}^m$, meaning that vector $\boldsymbol{a}$ has n real elements and vector $\boldsymbol{b}$ has m complex elements.

- A matrix $\boldsymbol{A}$ is called *square* if it has the same number of rows and columns: $\boldsymbol{A} = \boldsymbol{A}_{n\times n} = \boldsymbol{A}_n$. The elements $a_{11}, a_{22}, \ldots, a_{nn}$ of a square matrix $\boldsymbol{A}$ are said to be on the *main diagonal.*

- A square matrix $\boldsymbol{A}$ is called *upper triangular* if all its elements below the main diagonal are zero and *lower triangular* is all its elements above the main diagonal are zero. Both upper and lower triangular matrices are often called *triangular.*

- A square matrix $\boldsymbol{A}$ is called *diagonal* if all its elements outside of the main diagonal are zero.

- A diagonal matrix is called an *identity matrix* if all its diagonal elements are ones. The identity matrix is usually denoted as $\boldsymbol{I}$ or $\boldsymbol{I}_n$ to specify the dimension of the matrix:

$$\boldsymbol{I}_n = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix}.$$

Matrices can be added, multiplied by constants and by matrices. Here are the rules:

- *Multiplication by a constant.* Assume that we have a matrix $\boldsymbol{A}_{m\times n}$. To multiply this matrix by a constant α means to multiply every element of $\boldsymbol{A}$ by α, formally:

$$\boldsymbol{B}_{m\times n} = \alpha \boldsymbol{A}_{m\times n}, \quad b_{ij} = \alpha a_{ij}.$$

- *Addition of matrices.* To add two matrices we need to require these matrices have the same dimensions, then addition goes elementwise:

$$\boldsymbol{C}_{m\times n} = \boldsymbol{A}_{m\times n} + \boldsymbol{B}_{m\times n}, \quad c_{ij} = a_{ij} + b_{ij}.$$

  The rules of multiplication by constants and additions allow to make sense of, e.g., the following expression:

$$5\begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 2 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 2 & 3 \end{bmatrix} = 5\begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 2 & 1 \end{bmatrix} + (-1)\begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 & -5 \\ 10 & 0 \\ 10 & 5 \end{bmatrix} + \begin{bmatrix} -2 & 0 \\ 1 & -1 \\ -2 & -3 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ 11 & -1 \\ 8 & 2 \end{bmatrix}.$$

- *Matrix multiplication.* Two matrices $\boldsymbol{A}_{m\times k}$ and $\boldsymbol{B}_{k\times n}$ can be multiplied if and only if the number of columns of the first matrix is equal to the number of rows of the second matrix. The result will be matrix $\boldsymbol{C}_{m\times n}$:

  $$\boldsymbol{C}_{m\times n} = \boldsymbol{A}_{m\times k} \cdot \boldsymbol{B}_{k\times n}.$$

  The elements of the product can be found as

  $$c_{ij} = \sum_{l=1}^{k} a_{il} b_{lj}, \quad i = 1, \dots m, j = 1, \dots, n.$$

  The actual strategy to multiply matrices can be most easily learned by an example:

  $$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}_{3\times 2} \begin{bmatrix} -1 & -3 \\ -2 & -4 \end{bmatrix}_{2\times 2} = \begin{bmatrix} 1\cdot(-1)+2\cdot(-2) & 1\cdot(-3)+2\cdot(-4) \\ 3\cdot(-1)+4\cdot(-2) & 3\cdot(-3)+4\cdot(-4) \\ 5\cdot(-1)+6\cdot(-2) & 5\cdot(-3)+6\cdot(-4) \end{bmatrix}_{3\times 2}$$

  From the definition of matrix multiplication it is clear that even if we can multiply matrices $\boldsymbol{AB}$, it does not mean that $\boldsymbol{BA}$ makes sense. However, if $\boldsymbol{A}, \boldsymbol{B}$ are square matrices of the same *order* (the same dimension), then both $\boldsymbol{AB}$ and $\boldsymbol{BA}$ are defined and result in the matrices $n \times n$. Simple examples (provide one!) show that starting with dimension 2

  $$\boldsymbol{AB} \neq \boldsymbol{BA},$$

  i.e., matrix multiplication is not commutative.

  For any square matrix $\boldsymbol{A}_{n\times n}$ we have

  $$\boldsymbol{A}_{n\times n}\boldsymbol{I}_n = \boldsymbol{I}_n\boldsymbol{A}_{n\times n} = \boldsymbol{A}_{n\times n},$$

  which explains the name "identity matrix": It plays the same role as played by number one for number multiplication.

Putting together matrix addition and matrix multiplication, we can see now that indeed a system of linear first order ODE can be written in the concise matrix form

$$\dot{\boldsymbol{y}} = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{f}(t),$$

for the matrix $\boldsymbol{A}$, vector function $\boldsymbol{f} \colon \mathbf{R} \to \mathbf{R}^n$, and the unknown vector-function $\boldsymbol{y}(t)$. Especially handy here is the strange formula for matrix multiplication defined above, but of course the actual reason for this (not very natural and intuitive) definition is quite different from simply have a nice shortcut notation for systems of equations (including algebraic systems). You should definitely consider taking an intermediate level Linear Algebra course to find out why the matrix multiplication is defined in this way.

19.2 Determinants

Consider a system of two linear algebraic equations with two unknowns in the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2, \end{aligned}$$

or, in the matrix form

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

Multiply the first equation by a_{22}, second by a_{12}, and deduct the second one from the first. You'll find that

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12},$$

or, assuming that $a_{11}a_{22} - a_{12}a_{21} \neq 0$,

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}}.$$

Similarly, you can find for x_2:

$$x_2 = \frac{b_2a_{11} - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

These formulas provide a unique solution to our system provided, again, if $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Hence there is a quantity that is defined in terms of the elements of matrix $\boldsymbol{A}$, that indicates when our system has unique solution. This quantity is called the *determinant* of matrix $\boldsymbol{A}$ and denoted as $\det \boldsymbol{A}$ or $|\boldsymbol{A}|$. Hence we have

$$\det \boldsymbol{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

To define the determinant for a general square matrix $\boldsymbol{A}$ of the n-th order, I choose the recursive definition, which is convenient for determinant evaluations. In the usual linear algebra courses this definition follows from other, more conceptual definitions of determinants. First I will need the notion of a minor. The *minor* M_{ij} of the element a_{ij} in the matrix $\boldsymbol{A}$ is the determinant of the matrix, which is obtained from A by deleting the i-th row and the j-th column (hence the dimension reduces by one). The *cofactor* C_{ij} of the element a_{ij} is defined as $C_{ij} = (-1)^{i+j}M_{ij}$.

Definition 1. *The determinant of square matrix* $\boldsymbol{A}$ *is the number, which can be calculated by one of the following formulas:*

$$\det \boldsymbol{A} = \sum_{j=1}^{n} a_{ij}C_{ij},$$

for any $1 \leq i \leq n$*, or*

$$\det \boldsymbol{A} = \sum_{i=1}^{n} a_{ij}C_{ij},$$

for any $1 \leq j \leq n$*.*

Note that there are total $2n$ formulas in the above definition, and it is implicitly assumed that all of them yield the same answer (this is proved in the course of linear algebra). The first formula expands the determinant along row i, and the second formula expands the determinant along column j. Since each determinant is defined in terms of cofactors, i.e., the determinants of the matrices of size $n-1 \times n-1$, hence the definition is recursive. To finalize it we need, e.g., the determinant of matrix 2×2 (the formula is above), or we can define the determinant of matrix 1×1 to be equal the only element of this matrix: $\det[a_{11}] = a_{11}$.

Example 2. Using the definition, we can find the determinant of a 3×3 matrix, using the expansion along, e.g., the first row

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

The determinant is a function defined on the space of all square matrices and it possesses a number of useful to remember properties, which I list without proof:

- The determinant of a triangular matrix can be found as the product of the elements on the main diagonal:

  $$\det \boldsymbol{A} = a_{11} \dots a_{nn},$$

  if $\boldsymbol{A}$ is triangular. (*Q:* Can you prove this fact using the definition above?) As a simple corollary, we have

  $$\det \boldsymbol{I}_n = 1.$$

- If two rows (or two columns) are switched, then the determinant changes its sign.
- If a matrix has two identical rows or columns then its determinant is zero.
- If a matrix has a row of zeros, then its determinant is zero.
- $$\det \boldsymbol{A} = \det \boldsymbol{A}^\top.$$
- $$\det \alpha \boldsymbol{A} = \alpha^n \det \boldsymbol{A}.$$
- Adding a row multiplied by a constant to another row of a matrix does not change the determinant.
- $$\det \boldsymbol{AB} = \det \boldsymbol{A} \det \boldsymbol{B}.$$

While the definition for the determinant is good for evaluating the determinants of matrices of a reasonable size, in numerical computations usually the properties listed above are used to first put the matrix into triangular form with the transformations that do not change the determinant, and then calculate the determinant of the resulting triangular matrix.

The determinant gives a simple criterion when system of linear algebraic equations has a unique solution. Namely, system

$$\boldsymbol{Ax} = \boldsymbol{b},$$

of n equations with n unknowns has unique solution if and only if

$$\det \boldsymbol{A} \neq 0.$$

19.3 Inverse matrix and solving $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

Definition 3. *A matrix $\boldsymbol{B}$ of dimensions $n \times n$ is called the inverse to matrix $\boldsymbol{A}$ of dimensions $n \times n$ and denoted $\boldsymbol{A}^{-1}$ if*

$$\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_n.$$

The inverse, when it exists, is unique. To actually calculate the inverse matrix, I present an explicit formula. It should be noted, however, that this formula requires a significant amount of calculations and should *not* be used for matrices of order 5 and above.

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}}\boldsymbol{C}^{\top},$$

where

$$\boldsymbol{C} = [C_{ij}]_{n\times n},$$

is the matrix composed of cofactors of $\boldsymbol{A}$. This formula actually shows, among other things, that matrix $\boldsymbol{A}$ is *invertible* (i.e., it has the inverse) if and only if $\det \boldsymbol{A} \neq 0$. Such matrices are called *non-singular.*

Example 4. Find the inverse for 2×2 matrix. Let

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

We have

$$\det \boldsymbol{A} = ad - bc,$$

and

$$\boldsymbol{C} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}.$$

Therefore,

$$\boldsymbol{A}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Check that $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_2$.

Using the notion of the inverse matrix we can "solve" system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ in one line:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Leftrightarrow \boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} \Leftrightarrow \boldsymbol{I}_n\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} \Leftrightarrow \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}.$$

Actually, if you consider in details what is written in $\boldsymbol{A}^{-1}\boldsymbol{b}$, you will recover familiar *Cramer's formulas* for the solution of linear system:

$$x_i = \frac{\det \Delta_i}{\det \boldsymbol{A}}, \quad i = 1, \ldots, n,$$

where Δ_i is the matrix which is obtained from $\boldsymbol{A}$ by replacing the i-th column with column-vector $\boldsymbol{b}$.

Using the properties of the determinant and the formula for the inverse matrix, we have

$$\det \boldsymbol{A} = \frac{1}{\det \boldsymbol{A}^{-1}}.$$

Example 5. Solve system of three linear equations with three unknowns by the inverse matrix method:

$$\begin{aligned} x_1 + \quad +2x_3 &= 5, \\ x_1 - x_2 + x_3 &= 5, \\ 2x_1 + x_2 + x_3 &= 2. \end{aligned}$$

In matrix form we have

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \\ 2 & 1 & 1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 5 \\ 5 \\ 2 \end{bmatrix}.$$

First we find $\det \boldsymbol{A}$:

$$\det \boldsymbol{A} = 1 \cdot \begin{vmatrix} -1 & 1 \\ 1 & 1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} = -2 + 6 = 4.$$

The cofactor matrix is given here as

$$\boldsymbol{C} = \begin{bmatrix} -2 & 1 & 3 \\ 2 & -3 & -1 \\ 2 & 1 & -1 \end{bmatrix}.$$

Therefore, the inverse matrix is

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \boldsymbol{C}^{\top} = \frac{1}{4} \begin{bmatrix} -2 & 2 & 2 \\ 1 & -3 & 1 \\ 3 & -1 & -1 \end{bmatrix}.$$

You should check that $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$.

Finally, to solve the system, we need to evaluate

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \boldsymbol{A}^{-1}\boldsymbol{b} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}.$$

You can check the final results by plugging these values into the system.

To conclude this short review of the matrix arithmetics, I would like to note that there are much better ways to solve systems of linear algebraic equations than by finding the inverse matrix or using Cramer's rule. We will not need it in our course, but the right way to take care of actual computations if the *Gaussian elimination.*

19.4 Vector spaces and bases

Consider several examples of familiar objects in mathematics:

- The set of column-vectors in $\mathbf{R}^3$. I.e., these are the vectors of the form

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

  According to the discussed rules, we can add two vectors, and we can multiply a vector by a real number. The result is again a vector in $\mathbf{R}^3$.

- The set of column-vectors in $\mathbf{C}^n$. This is not really different from the previous example, however now we can multiply by complex constants.

- The set of real matrices of dimension $m \times n$, let us denote this set as $\mathbf{M}_{m\times n}$. We know that we can add matrices of the same dimension and multiply them by a real constant. The result will be in $\mathbf{M}_{m\times n}$.

- Consider the fourth order linear homogeneous ODE with constant coefficients:

$$y^{(4)} + a_3 y''' + a_2 y'' + a_1 y' + a_0 y = 0.$$

  Recall from the previous part of the course that if I denote the set of all the solutions to this equation as $\mathcal{S}$, then if $y_1, y_2 \in \mathcal{S}$ (i.e., y_1 and y_2 are solutions) then their linear combination $\alpha y_1 + \beta y_2$ is also a solution for any constants $\alpha, \beta \in \mathbf{R}$.

- Consider a homogeneous system of linear algebraic equations:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0},$$

  where $\boldsymbol{A} = [a_{ij}]_{m\times n}$. This system always has at least one solution (the trivial one). Denote the set of all solutions as $\mathcal{X}$. Recall from your linear algebra course (or prove by yourself), that if $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{X}$ then $\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2 \in \mathcal{X}$ for any $\alpha, \beta \in \mathbf{R}$.

- Let $\mathcal{P}_n$ denote the set of all polynomials of degree at most n with real coefficients. Then if you take two polynomials $P_1, P_2 \in \mathcal{P}_n$ then their linear combination $\alpha P_1 + \beta P_2 \in \mathcal{P}_n$.

What is common between sets $\mathbf{R}^3, \mathbf{C}^n, \mathbf{M}_{m\times n}, \mathcal{S}, \mathcal{X}, \mathcal{P}_n$ in the above examples? For all of them it is true that any linear combination of the elements of these sets still belongs to the same set. The mathematical abstraction that deals with such structures is called the *vector* (or *linear*) *space*. Hence $\mathbf{R}^3, \mathbf{C}^n, \mathbf{M}_{m\times n}, \mathcal{S}, \mathcal{X}, \mathcal{P}_n$ are examples of vector spaces. Here is the formal definition.

Definition 6. *A vector space $\mathcal{V}$ over the real numbers $\mathbf{R}$ (or over the complex numbers $\mathbf{C}$) is a nonempty set with the operations of addition, such that for any two elements $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathcal{V} \implies \boldsymbol{v}_1 + \boldsymbol{v}_2 \in \mathcal{V}$, and multiplication by real scalars (complex scalars): $\alpha\boldsymbol{v} \in \mathcal{V}$ for any $\alpha \in \mathbf{R}$ ($\alpha \in \mathbf{C}$) and any $\boldsymbol{v} \in \mathcal{V}$. For these two operations the following axioms hold:*

$$\begin{aligned}
\boldsymbol{v}_1 + \boldsymbol{v}_2 &= \boldsymbol{v}_2 + \boldsymbol{v}_1, \\
(\boldsymbol{v}_1 + \boldsymbol{v}_2) + \boldsymbol{v}_3 &= \boldsymbol{v}_1 + (\boldsymbol{v}_2 + \boldsymbol{v}_3), \\
\textit{there exists } 0 \in \mathcal{V},\ 0 + \boldsymbol{v} &= \boldsymbol{v}, \\
\textit{there exists } -\boldsymbol{v},\ \boldsymbol{v} + (-\boldsymbol{v}) &= 0, \\
\alpha(\beta\boldsymbol{v}) &= (\alpha\beta)\boldsymbol{v}, \\
1 \cdot \boldsymbol{v} &= \boldsymbol{v}, \\
\alpha(\boldsymbol{v}_1 + \boldsymbol{v}_2) &= \alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2, \\
(\alpha + \beta)\boldsymbol{v} &= \alpha\boldsymbol{v} + \beta\boldsymbol{v}.
\end{aligned}$$

Here $\alpha, \beta \in \mathbf{R}$ (or in $\mathbf{C}$), and $\boldsymbol{v}, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \in \mathcal{V}$. The elements of a vector space are called vectors.

It is a good exercise to convince yourself that for all the examples above, the sets $\mathbf{R}^3, \mathbf{C}^n, \mathbf{M}_{m\times n}, \mathcal{S}$, $\mathcal{X}, \mathcal{P}_n$ are vector spaces, for which all the listed axioms hold. Hence in each case the elements of these sets are *vectors*. This is a confusing point, since we usually get used to the point that vectors are arrays of numbers as in the first two examples. However, a matrix, a solution to the linear differential equation, and a polynomial is also a vector. There is a good reason to call arrays of numbers as vectors, but for these details you need to consult your linear algebra course.

Why this abstract definition of a vector space useful? Because in many cases we can express any element of the vector space using its *basis*. Here are some more definitions.

Consider an ordered set of vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$. This set is said to be *linearly independent* if any linear combination of these vectors

$$\alpha_1 \boldsymbol{v}_1 + \ldots \alpha_k \boldsymbol{v}_k$$

is equal to zero only if $\alpha_1 = \ldots = \alpha_k = 0$. This ordered set of vectors is said to be *linearly dependent* if there exist constants $\alpha_1, \ldots, \alpha_k$ not equal to zero simultaneously such that

$$\alpha_1 \boldsymbol{v}_1 + \ldots \alpha_k \boldsymbol{v}_k = 0.$$

With a slight abuse of the language I will call the vectors belonging to a linearly independent (or dependent) set as simply linearly dependent (or independent).

The notion of the linear independence means that none of the vectors in the set can be expressed as a linear combination of the rest of them; if vectors are linearly dependent, then there is linear combination expressing at least one of the vectors through the rest.

The *span* of the set of vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is by definition all possible linear combinations of these vectors. Vector space $\mathcal{V}$ is called *finite dimensional* if there exits a finite set of vectors that span $\mathcal{V}$, otherwise $\mathcal{V}$ is called *infinite dimensional*. A *basis* of vector space $\mathcal{V}$ is a set of vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ that 1) is linearly independent and 2) spans $\mathcal{V}$. The *dimension* of a finite-dimensional vector space $\mathcal{V}$ is the number of vectors in a basis. This is a basic fact from linear algebra that any basis in finite dimensional vector space has the same number of elements, and hence the definition of the dimension makes perfect sense.

Example 7. Vector space $\mathbf{R}^3$ is 3-dimensional. To show this fact we need to present a basis. The standard basis of $\mathbf{R}^3$ is

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Why is this set of vectors a basis? Because, first, any vector is in span these vectors, i.e., any vector can be presented as a linear combination:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \boldsymbol{e}_1 + x_2 \boldsymbol{e}_2 + x_3 \boldsymbol{e}_3.$$

It is said that x_1, x_2, x_3 are the coordinates of $\boldsymbol{x}$ in the basis $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$. Second, we need to show that these three vectors are linearly independent. Assume opposite: Let $\alpha_1, \alpha_2, \alpha_3$ be constants not equal to zero simultaneously, such that

$$\alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2 + \alpha_3 \boldsymbol{e}_3 = 0.$$

The last expression can be rewritten in the matrix form

$$\boldsymbol{I}\boldsymbol{\alpha} = 0, \quad \boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}.$$

The vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ are the columns of the matrix $\boldsymbol{I}$. This is a system of 3 linear algebraic homogeneous equations with 3 unknowns. Since $\det \boldsymbol{I} = 1$, it has a unique solution, which is zero vector. Therefore $\boldsymbol{\alpha} = 0$. Contradiction.

Actually, it is possible to generalize this example in the following way: in $\mathbf{R}^n$ n vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ form a basis if and only if the matrix $\boldsymbol{V}$, where the i-th column is given by $\boldsymbol{v}_i$, has nonzero determinant:

$$\det \boldsymbol{V} \neq 0.$$

As a corollary, we have that $n+1$ vectors in $\mathbf{R}^n$ are always linearly dependent.

Example 8. Vector space of solutions $\mathcal{S}$ to the equation

$$y'' + y = 0$$

is two dimensional. To show this fact we need to present a basis. I claim that the basis can be taken as $\{\cos t, \sin t\}$. The general solution to this equation has the form

$$y(t) = C_1 \cos t + C_2 \sin t,$$

therefore any solution can be represented as a linear combination of cos and sin. Moreover, these two functions are linearly independent on any interval I. To prove this we can use the Wronskian, or we can do it directly, by considering the linear combination

$$\alpha \cos t + \beta \sin t = 0.$$

Note that expression on the left can be rearranged as

$$\sqrt{\alpha^2 + \beta^2} \cos(t + \varphi) = 0,$$

and this expression equals to zero identically on any interval if and only if $\sqrt{\alpha^2 + \beta^2} = 0$, which means that $\alpha = \beta = 0$, which proves linear independence.

Assume that we also given initial conditions $y(0) = 1$, $y'(0) = 0$. Then the solution is

$$y(t) = 1 \cdot \cos t + 0 \cdot \sin t = \cos t.$$

It is said that this solution has coordinates $(1, 0)$ in the basis $\{\cos t, \sin t\}$.