8 An introduction to linear algebra

(Applied Mathematics — FAPPZ)

Petr Gurka

(Updated on November 28, 2011)

Contents

1 Vectors, vector spaces

1.1 Motivation

Consider a system of linear equations

$$\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 1 \\ 2x & + & 5y & & & = & 9 \\ 3x & + & 6y & - & 2z & = & 6 \end{array} . \tag{1}$$

To solve the equation (1) means to find numbers x, y, z which satisfy all the equations of the system.

How to expres the system using matrices or vectors

The system can be expressed in matrix form:

$$\left(\begin{array}{rrr|r} 1 & 2 & -1 & 1 \\ 2 & 5 & 0 & 9 \\ 3 & 6 & -2 & 6 \end{array}\right). \tag{2}$$

The matrix (2) is called the *augmented matrix* of the system (1).

Notice that the system (1) can be also expressed using a "linear combination" of "arithmetic vectors", that is,

$$x\begin{pmatrix}1\\2\\3\end{pmatrix}+y\begin{pmatrix}2\\5\\6\end{pmatrix}+z\begin{pmatrix}-1\\0\\-2\end{pmatrix}=\begin{pmatrix}1\\9\\6\end{pmatrix}.$$

1.2 Arithmetic vector space

Definition 1 (Arithmetic vector space)**.** By $\mathbf{V}_n$, $n \in \mathbb{N}$, we denote the *arithmetic vector space.* It consists of ordered n-tuples of real numbers, i. e.

$$\mathbf{V}_n = \{(a_1, \ldots, a_n) \mid a_1, \ldots, a_n \in \mathbb{R}\}.$$

Elements of $\mathbf{V}_n$ we call vectors. We define a sum of two vectors and multiplication of a vector by a (real) number:

sum	$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$,
multiplication	$r\,(a_1, \ldots, a_n) = (r\,a_1, \ldots, r\,a_n)$.

We denote vectors by bold letters as $\mathbf{a}$, $\mathbf{b}$ etc.

The vector $\mathbf{o} = (0, \ldots, 0)$ we call a *zero vector,* a vector $-\mathbf{a} = (-a_1, \ldots, -a_n)$ we call the *opposite vector* to $\mathbf{a} = (a_1, \ldots, a_n)$.

1.3 Vector space

We can define a vector space in a more general sense.

Definition 2 (Vector space)**.** Nonempty set $\mathbf{V}$ with operations $\mathbf{u} + \mathbf{v} \in \mathbf{V}$, a *sum* of elements $\mathbf{u}, \mathbf{v} \in \mathbf{V}$, and $r\,\mathbf{u} \in \mathbf{V}$, a *multiplication* of $\mathbf{u} \in \mathbf{V}$ by a real number r, we call the vector space if, for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{V}$, and $r, s \in \mathbb{R}$, the following conditions are satisfied:

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
2. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$;
3. there exists $\mathbf{o} \in \mathbf{V}$ such that $\mathbf{u} + \mathbf{o} = \mathbf{u}$;
4. $r(\mathbf{u} + \mathbf{v}) = r\,\mathbf{u} + r\,\mathbf{v}$;
5. $(r + s)\mathbf{u} = r\,\mathbf{u} + s\,\mathbf{u}$;
6. $r(s\,\mathbf{u}) = (r\,s)\mathbf{u}$;
7. $1\,\mathbf{u} = \mathbf{u}$, $0\,\mathbf{u} = \mathbf{o}$.

Elements of $\mathbf{V}$ we call *vectors*, the element $\mathbf{o}$ we call the *zero vector.*

1.4 Linear combination

Definition 3. Let $\mathbf{u}, \mathbf{u}_1, \ldots, \mathbf{u}_k$ be vectors, $r_1, \ldots, r_k$ real numbers and

$$\mathbf{u} = r_1\,\mathbf{u}_1 + \cdots + r_k\,\mathbf{u}_k.$$

Then we say that the *vector* $\mathbf{u}$ is a *linear combination* of vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ with coefficients $r_1, \ldots, r_k$.

1.5 Linear independence and dependence

Definition 4. We say that vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ are *linearly independent* if the following condition holds:

$$\text{whenever} \quad r_1\,\mathbf{u}_1 + \cdots + r_k\,\mathbf{u}_k = \mathbf{o}, \quad \text{then} \quad r_1 = r_2 = \cdots = r_k = 0.$$

If the vectors are not linearly independent we call them *linearly dependent*.

Remark 5.

- If one of the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ is a zero vector then they are linearly *dependent*.
- One vector $\mathbf{u}$ is linearly *dependent* if and only if it is a *zero vector*.
- Two vectors $\mathbf{u}$, $\mathbf{v}$ are linearly *dependent* if and only if one of them is a constant multiple of the other.

1.6 Linear span

Definition 6. The set of all linear combinations of given vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$, that is, the set of all $\mathbf{u} = r_1\,\mathbf{u}_1 + \cdots + r_k\,\mathbf{u}_k$, we call a *linear span* of vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$. We denote it by $\langle \mathbf{u}_1, \ldots, \mathbf{u}_k \rangle$. Vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$ we call *generators* of the linear span $\langle \mathbf{u}_1, \ldots, \mathbf{u}_k \rangle$. We also say that *vectors* $\mathbf{u}_1, \ldots, \mathbf{u}_k$ *generate* the linear span $\langle \mathbf{u}_1, \ldots, \mathbf{u}_k \rangle$.

2 Elementary operations with vectors

2.1 Collection of vectors

Definition 7. By a *finite collection of vectors* $[\mathbf{u}_1, \ldots, \mathbf{u}_k]$ we mean vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k$, where some of them may be equal.

In contrast to a set, a collection can contain the same element more times. For example, from the set of vectors $\{\mathbf{a}, \mathbf{b}\}$ we can make collections $[\mathbf{a}, \mathbf{b}]$, $[\mathbf{a}, \mathbf{a}, \mathbf{b}]$, $[\mathbf{a}, \mathbf{a}, \mathbf{b}, \mathbf{a}, \mathbf{b}, \mathbf{b}]$. From the collection of vectors we can select the set which contains all its vectors, e. g. from the collection $[\mathbf{a}, \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{c}]$ we select the set $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$.

2.2 Elementary operations

Definition 8. Under an *elementary operation* with a collection $[\mathbf{u}_1, \ldots, \mathbf{u}_k]$ we understand one of the following operations:

1. interchanging vectors in the collection;
2. multiplying arbitrary vector by a nonzero constant;
3. adding a constant multiple of a vector to another vector;
4. removing a zero vector from the collection, provided it is not the only vector in the collection;
5. removing a vector from the collection whenever the collection contains another vector which is its multiple.

Theorem 9. *Applying any of elementary operations* (1)–(5) *to a collection of vectors we obtain a collection which generates the same linear span as the original collection.*

It is useful to introduce matrices.

Definition 10. A matrix $\boldsymbol{A}$ of type (m, n) we call a rectangular array of mn real numbers in m rows and n columns,

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

A matrix $\boldsymbol{A}$ is called a *zero matrix* if it contains only zero entries, that is, $a_{ij} = 0$, $i = 1, \dots, m$, $j = 1, \dots, n$. We denote a zero matrix by $\boldsymbol{O}$.

2.3 Equivalent matrices

Definition 11. Two matrices $\boldsymbol{A}$, $\boldsymbol{B}$ are called equivalent if the linear span generated by rows of the matrix $\boldsymbol{A}$ is the same as the linear span generated by rows of the matrix $\boldsymbol{B}$. Then we write $\boldsymbol{A} \sim \boldsymbol{B}$.

Remark 12. It immediately follows from Theorem 9 that applying any of elementary operations (1)–(5) to rows of a matrix $\boldsymbol{A}$ we obtain a matrix which is equivalent to $\boldsymbol{A}$.

2.4 Gaussian matrix

Definition 13. We say that a nonzero matrix $\boldsymbol{B}$ is a Gaussian matrix if the first nonzero number of any of its rows (considered from left to right) is the last nonzero number of any of its columns (considered down from top). This number is called the *leading entry* of the corresponding row.

Remark 14. A Gaussian matrix is also called a matrix in row-echelon form.

Example 15. The following two matrices are Gaussian matrices

$$\begin{pmatrix} 1 & -2 & 0 & 4 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & 0 & 5 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The matrix

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

is not a Gaussian matrix.

2.5 Rank of the matrix

Definition 16. Under the *rank* of a nonzero matrix $\boldsymbol{A}$ we understand the maximal number of the linearly independent rows of the matrix $\boldsymbol{A}$ which generate the same linear span as all rows of the matrix $\boldsymbol{A}$. We denote it by $\mathrm{h}(\boldsymbol{A})$. For a zero matrix we put $\mathrm{h}(\boldsymbol{O}) = 0$

Theorem 17. *A Gaussian matrix $\boldsymbol{B}$ has linearly independent rows. Consequently, $\mathrm{h}(\boldsymbol{B})$ is equal to the number of its rows.*

Theorem 18. *Let $\boldsymbol{A}$ be a nonzero matrix. Then there exists a Gaussian matrix $\boldsymbol{B}$ such that $\boldsymbol{A} \sim \boldsymbol{B}$.*

Proof. The assertion follows by the Gaussian elimination which will be described later. □

Theorem 19. *Let $\boldsymbol{A}$ be a nonzero matrix. Let $\boldsymbol{B}_1$, $\boldsymbol{B}_2$ be two Gaussian matrices such that $\boldsymbol{A} \sim \boldsymbol{B}_1$ a $\boldsymbol{A} \sim \boldsymbol{B}_2$. Then both matrices $\boldsymbol{B}_1$, $\boldsymbol{B}_2$ have the same number of rows which is equal to $\mathrm{h}(\boldsymbol{A})$.*

2.6 Gaussian elimination

Method of Gaussian elimination (in this context) is based on clever application of elementary operations (1)–(5) with rows of the matrix $\boldsymbol{A}$.

Here is a systematic way how to do it:

1. Start by obtaining 1 in the top left corner. Then obtain zeros below that 1 by adding appropriate multiples of the first row to the rows below it.
2. Next, obtain a leading 1 in the next row, and then obtain zeros below that 1.
3. At each stage make sure that every leading entry is to the right of the leading entry in the row above it — rearrange the rows if necessary.
4. Continue this process until you arrive at a Gaussian matrix.

(Instead of the number 1 we can have any nonzero number.)