eNote 11

General Vector Spaces

In this eNote a general theory is presented for all mathematical sets where addition and multiplication by a scalar are defined and which satisfy the same arithmetic rules as geometric vectors in the plane and in 3-space. Using the concepts of a basis and coordinates, it is shown how one can simplify and standardize the solution of problems that are common to all these sets, which are called vector spaces. Knowledge of eNote 10 about geometric vectors is an advantage as is knowledge about the solution sets for systems of linear equations, see eNote 6. Finally, elementary matrix algebra and a couple of important results about determinants are required (see eNotes 7 and 8).

Updated: 4.10.21 David Brander

11.1 Generalization of the Concept of a Vector

The vector concept originates in the geometry of the plane and space where it denotes a pair consisting of a length and a direction. Vectors can be represented by a line segment with orientation (an arrow) following which it is possible to define two geometric operations: *addition* of vectors and *multiplication* of vectors *by numbers* (scalar). For the use in more complicated arithmetic operations one proves eight arithmetic rules concerning the two arithmetic operations.

In many other sets of mathematical objects one has a need for defining addition of the objects and multiplication of an object by a scalar. The number spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ and the set of matrices $\mathbb{R}^{m\times n}$ are good examples, see eNote 5 and eNote 6, respectively. The remarkable thing is, that the *arithmetic rules* for addition and multiplication by a scalar, that are possible to prove within each of these sets, are the same as the arithmetic rules for geometric vectors in the plane and in space! Therefore one says: Let us make a *theory* that applies to all the sets where addition and multiplication by a scalar can be defined and where all the eight arithmetic rules known from geometry are valid. By this one carries out a *generalization* of the concept of geometric vectors, and every set that obeys the conditions of the theory is therefore called a *vector space*.

In eNote 10 about geometric vectors it is demonstrated how one can introduce a *basis* for the vectors following which the vectors are determined by their *coordinates* with respect to this basis. The advantage of this is that one can replace the geometric vector calculation by calculation with the coordinates for the vector. It turns out that it is also possible to transfer the concepts of basis and coordinates to many other sets of mathematical objects that have addition and multiplication by a scalar.

In the following, when we investigate vector spaces in the abstract sense, it means that we look at which concepts, theorems and methods follow from the common arithmetic rules, as we ignore the concrete meaning of addition and multiplication by a scalar has in the sets of concrete objects where they are introduced. By this one obtains general methods for *every* set of the kind described above. The application in any particular vector space then calls for *interpretation* in the context of the results obtained. The approach is called *the axiomatic method*. Concerning all this we now give the abstract definition of vector spaces.

|||| Definition 11.1 Vector Spaces

Let $\mathbb{L}$ denote $\mathbb{R}$ or $\mathbb{C}$, and let V be a set of mathematical elements where there is defined two arithmetic operations:

I. *Addition* that from two elements $\mathbf{a}$ and $\mathbf{b}$ in V forms the sum $\mathbf{a}+\mathbf{b}$ that also belongs to V.

II. *Multiplication by a scalar* that from any $\mathbf{a} \in V$ and any scalar $k \in \mathbb{L}$ forms a product $k\mathbf{a}$ or $\mathbf{a}k$ that also belongs to V.

V is called a ***vector space*** and the elements of V ***vectors*** if the following eight arithmetic rules are valid:

1.	$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$	Addition is commutative
2.	$(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$	Addition is associative
3.	$\mathbf{a}+\mathbf{0}=\mathbf{a}$	In V there exists $\mathbf{0}$ that is *neutral* wrt. addition
4.	$\mathbf{a}+(-\mathbf{a})=\mathbf{0}$	For every $\mathbf{a} \in V$ there is an *opposite object* $-\mathbf{a} \in V$
5.	$k_1(k_2\mathbf{a})=(k_1k_2)\mathbf{a}$	Product by a scalar is associative
6.	$(k_1+k_2)\mathbf{a}=k_1\mathbf{a}+k_2\mathbf{a}$	} The distributive rule applies
7.	$k_1(\mathbf{a}+\mathbf{b})=k_1\mathbf{a}+k_1\mathbf{b}$	
8.	$1\mathbf{a}=\mathbf{a}$	The scalar 1 is *neutral* in products with vectors

If $\mathbb{L}$ in the definition 11.1 stands for $\mathbb{R}$ then V is a ***vector space over the real numbers***. This means that the scalar k (only) can be an arbitrary real number. Similarly one talks about V as a ***vector space over the complex numbers*** if $\mathbb{L}$ stands for $\mathbb{C}$, where k can be an arbitrary complex number.

The requirements I and II in the definition 11.1, that the results of addition and of multiplication by a scalar in itself must be an element in V, are called the ***stability requirements***. In other words V must be stable with respect to the two arithmetic operations.

The set of geometric vectors in the plane and the set of geometric vectors in space are naturally the most obvious examples of vector spaces, since the eight arithmetic rules in the definition 11.1 are constructed from the corresponding rules for geometric vectors. But let us check the *stability requirements*. Is the sum of two vectors in the plane itself a vector in the plane? And is a vector in the plane multiplied by a number in itself a vector in the plane? From the definition of the two arithmetic operations (see Definition 10.2 and Definition 10.3), the answer is obviously yes, therefor the set of vectors in the plane is a vector space. Similarly we see that the set of vectors in 3-space is a vector space.

|||| Theorem 11.2 Uniqueness of the 0-Vector and the Opposite Vector

For every vector space V:

1. V only contains one neutral element with respect to addition.
2. Every vector $\mathbf{a} \in V$ has only one opposite element.

|||| Proof

First part:
Let $\mathbf{0}_1$ and $\mathbf{0}_2$ be two elements in V both neutral with respect to addition. Then:

$$\mathbf{0}_1 = \mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_2\,,$$

where we have used the fact that addition is commutative. There is only one 0-vector: $\mathbf{0}$.

Second part:
Let $\mathbf{a}_1, \mathbf{a}_2 \in V$ be two opposite elements for $\mathbf{a} \in V$. Then:

$$\mathbf{a}_1 = \mathbf{a}_1 + \mathbf{0} = \mathbf{a}_1 + (\mathbf{a} + \mathbf{a}_2) = (\mathbf{a} + \mathbf{a}_1) + \mathbf{a}_2 = \mathbf{0} + \mathbf{a}_2 = \mathbf{a}_2\,,$$

where we have used the fact that addition is both commutative and associative. Hence there is for $\mathbf{a}$ only one opposite vector $-\mathbf{a}$.

■

|||| Definition 11.3 Subtraction

Let V be a vector space, and let $\mathbf{a}, \mathbf{b} \in V$. By the difference $\mathbf{a} - \mathbf{b}$ we understand the vector

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})\,. \tag{11-1}$$

|||| Exercise 11.4

Prove that $(-1)\mathbf{a} = -\mathbf{a}$.

|||| Exercise 11.5 Zero-Rule

Prove that the following variant of the zero-rule applies to any vector space:

$$k\mathbf{a} = \mathbf{0} \Leftrightarrow k = 0 \;\text{ or }\; \mathbf{a} = \mathbf{0}\,. \tag{11-2}$$

|||| Example 11.6 Matrices as Vectors

For two arbitrary natural numbers m and n, $\mathbb{R}^{m\times n}$ (that is, the set of real $m \times n$-matrices) is a vector space. Similarly $\mathbb{C}^{m\times n}$ (that is, the set of complex $m \times n$-matrices) is a vector space

Consider e.g. $\mathbb{R}^{2\times 3}$. If we add two matrices of this type we get a new matrix of the same type, and if we multiply a 2×3-matrix by a number, we also get a new 2×3-matrix (see Definition 7.1). Thus the stability requirements are satisfied. That $\mathbb{R}^{2\times 3}$ in addition satisfies the eight arithmetic rules, is apparent from Theorem 7.3.

|||| Exercise 11.7

Explain that for every natural number n the number space $\mathbb{L}^n$ is a vector space. Remember to think about the case $n = 1$!

In the following two examples we shall see that the geometrically inspired vector space theory, surprisingly, can be applied to well known sets of *functions*. Mathematic historians have in this connection talked about *the geometrization of mathematical analysis*!

|||| Example 11.8 Polynomials as Vectors

The set of polynomials $P : \mathbb{R} \mapsto \mathbb{R}$ of at most n'th degree is denoted $P_n(\mathbb{R})$. An element P in $P_n(\mathbb{R})$ is given by

$$P(x) = a_0 + a_1 x + \cdots + a_n x^n \tag{11-3}$$

where the coefficients $a_0, a_1, \cdots a_n$ are arbitrary real numbers. Addition of two polynomials in $P_n(\mathbb{R})$ is defined by pairwise addition of coefficients belonging to the same degree of the variable, and multiplication of a polynomial in $P_n(\mathbb{R})$ by a number k is defined as the multiplication of every coefficient with k. As an example of the two arithmetic operations we look at two polynomials from $P_3(\mathbb{R})$:

$$P(x) = 1 - 2x + x^3 = 1 - 2x + 0x^2 + 1x^3$$

and

$$Q(x) = 2 + 2x - 4x^2 = 2 + 2x - 4x^2 + 0x^3.$$

By the sum of P and Q we understand the polynomial $R = P + Q$ given by

$$R(x) = (1+2) + (-2+2)x + (0-4)x^2 + (1+0)x^3 = 3 - 4x^2 + x^3$$

and by the multiplication of P by the scalar $k = 3$ we understand the polynomial $S = 3P$ given by

$$S(x) = (3 \cdot 1) + (3 \cdot (-2))x + (3 \cdot 0)x^2 + (3 \cdot 1)x^3 = 3 - 6x + 3x^3.$$

We will now justify that $P_n(\mathbb{R})$ with the introduced arithmetic operations is a vector space! That $P_n(\mathbb{R})$ satisfies the *stability requirements* follows from the fact that the sum of two polynomials of at most n'th degree in itself is a polynomial of at most n'th degree, and that multiplication of a polynomial of at most n'th degree by a real number again gives a polynomial of at most n'th degree. The conditions 1, 2 and 5 - 8 in the definition 11.1 are satisfied, because the same rules of operation apply to the calculations with coefficients of the polynomials that are used in the definition of the operations. Finally the conditions 3 and 4 are satisfied since the polynomial

$$P(x) = 0 + 0x + \cdots 0x^n = 0$$

constitutes the *zero vector*, and *the opposite vector* to $P(x) \in P_n(\mathbb{R})$ is given by the polynomial

$$-P(x) = -a_0 - a_1 x - \cdots - a_n x^n.$$

In the same way we show that polynomial $P : \mathbb{C} \mapsto \mathbb{C}$ of at most n'th degree, which we denote by $P_n(\mathbb{C})$, is a vector space.

|||| Exercise 11.9

Explain that $P(\mathbb{R})$, that is the set of real polynomials, is a vector space.

|||| Example 11.10 Continuous Functions as Vectors

The set of continuous real functions on a given interval $I \subseteq \mathbb{R}$ is denoted $C^0(I)$. Addition $m = f + g$ of two functions f and g in $C^0(I)$ is defined by

$$m(x) = (f+g)(x) = f(x) + g(x) \text{ for every } x \in I$$

and multiplication $n = k \cdot f$ of the function f by a real number k by

$$n(x) = (k \cdot f)(x) = k \cdot f(x) \text{ for every } x \in I\,.$$

We will now justify that $C^0(I)$, with the introduced operations of calculations, is a vector space. Since $f + g$ and $k \cdot f$ are continuous functions, we see that $C^0(I)$ satisfies the two stability requirements. Moreover: there exists a function that acts as the zero vector, viz. the zero function, that is, the function that has the value 0 for all $x \in I$, and the opposite vector to $f \in C^0(I)$ is the vector $(-1)f$ that is also written $-f$, and which for all $x \in I$ has the value $-f(x)$. Now it is obvious that $C^0(I)$ with the introduced operations of calculation satisfies all eight rules in definition 11.1, and $C^0(I)$ is thus a vector space.

11.2 Linear Combinations and Span

A consequence of arithmetic rules such as $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ and $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ from the definition 11.1 is that one can omit parentheses when one adds a series of vectors: the order of vector addition has no influence on the resulting sum vector. This is the background for *linear combinations* where a set of vectors is multiplied by a scalar and thereafter written as a sum.

|||| Definition 11.11 Linear Combination

When in a vector space V p vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ are given, and arbitrary scalars $k_1, k_2, \ldots, k_p$ are chosen, then the sum

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_p\mathbf{v}_p$$

is called a ***linear combination*** of the p given vectors.

If all the $k_1, \cdots, k_p$ are equal to 0, the linear combination is called *improper*, or *trivial*, but if at least one of the scalars is different from 0, it is called *proper* or *non-trivial*.

In the definition 11.11 only one linear combination is mentioned. In many circumstances it is of interest to consider the total set of possible linear combinations of given vectors. The set is called the *span* of the vectors. Consider e.g. a plane in space, through the origin and containing the position vectors for two non-parallel vectors **u** and **v**. The plane can be considered the span of the two vectors since the position vectors

$$\overrightarrow{OP} = k_1\mathbf{u} + k_2\mathbf{v}$$

"run through" all points P in the plane when k_1 and k_2 take on all conceivable real values, see Figure 11.1.

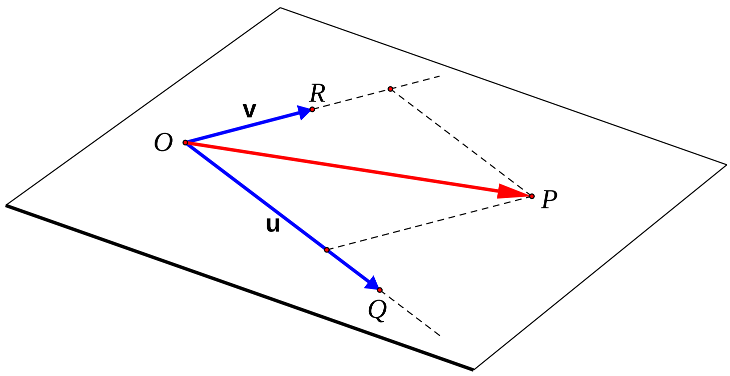

Figure 11.1: **u** and **v** span a plane in space

|||| Definition 11.12 Span

By the ***span*** of a given set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in a vector space V we understand the total set of all possible linear combinations of the vectors. The span of the p vectors is denoted by

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}\,.$$

|||| Example 11.13 Linear Combination and Span

We consider in the vector space $\mathbb{R}^{2\times 3}$ the three matrices/vectors

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & 2 \end{bmatrix},\ \mathbf{B} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} \text{ and } \mathbf{C} = \begin{bmatrix} -1 & -2 & 9 \\ 0 & 0 & 4 \end{bmatrix}. \tag{11-4}$$

An example of a linear combination of the three vectors is

$$2\mathbf{A} + 0\,\mathbf{B} + (-1)\mathbf{C} = \begin{bmatrix} 3 & 2 & -3 \\ 0 & 4 & 0 \end{bmatrix}. \tag{11-5}$$

We can then write

$$\begin{bmatrix} 3 & 2 & -3 \\ 0 & 4 & 0 \end{bmatrix} \in \operatorname{span}\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}\,. \tag{11-6}$$

11.3 Linear Dependence and Linear Independence

Two geometric vectors $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent if they are parallel, that is if there exists a number k, such that $\mathbf{v} = k\mathbf{u}$. More generally an arbitrary set of vectors are linearly dependent if one of the vectors is a linear combination of the others. We wish to transfer this concept to vector space theory:

|||| Definition 11.14 Linear Dependence and Independence

A set consisting of p vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ in a vector space V is ***linearly dependent*** if at least one of the vectors can be written as a linear combination of the others: for example

$$\mathbf{v}_1 = k_2\mathbf{v}_2 + k_3\mathbf{v}_3 + \cdots + k_p\mathbf{v}_p\,.$$

If none of the vectors can be written as a linear combination of the others, the set is called ***linearly independent***.

NB: If the set of vectors only consists of a single vector, the set is called linearly dependent if it consists of the 0-vector, and otherwise linearly independent.

|||| Example 11.15 Linear Dependence

Any set of vectors containing the zero vector, is linearly dependent! Consider e.g. the set $\{\mathbf{u}, \mathbf{v}, \mathbf{0}, \mathbf{w}\}$), here the zero vector can trivially be written as a linear combination of the three other vectors in the set:

$$\mathbf{0} = 0\mathbf{u} + 0\mathbf{v} + 0\mathbf{w}\,,$$

where the zero vector is written as a linear combination of the other vectors in the set.

|||| Example 11.16 Linear Dependence

Consider in the vector space $\mathbb{R}^{2\times 3}$ the three matrices/vectors

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & 2 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} \text{ and } \mathbf{C} = \begin{bmatrix} -1 & -2 & 9 \\ 0 & 0 & 4 \end{bmatrix}. \qquad (11\text{-}7)$$

$\mathbf{C}$ can be written as a linear combination of $\mathbf{A}$ and $\mathbf{B}$ since

$$\mathbf{C} = 3\mathbf{A} - 2\mathbf{B}\,.$$

Therefore $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are linearly dependent.

In contrast the set consisting of $\mathbf{A}$ and $\mathbf{B}$ is linearly independent, because these two vectors are not "parallel', since a number k obviously does not exist such that $\mathbf{B} = k\mathbf{A}$. Similarly with the sets $\{\mathbf{A}, \mathbf{C}\}$ and $\{\mathbf{B}, \mathbf{C}\}$.

When you investigate whether a set of vectors is linearly dependent, use of the definition 11.14 provokes the question *which* of the vectors is a linear combination of the others. Where should we begin the investigation? The dilemma can be avoided if bypassing the definition we instead use the following theorem:

|||| Theorem 11.17 Linear Dependence and Independence

A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ in a vector space V is linearly dependent if and only if the zero vector can be written as proper linear combination of the vectors – that is, if and only if scalars $k_1, k_2, \ldots, k_p$ exist that are not all equal to 0, and that satisfy

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_p\mathbf{v}_p = \mathbf{0}\,. \qquad (11\text{-}8)$$

. Otherwise the vectors are linearly independent.

.

|||| Proof

Assume first that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ are linearly dependent, then one can be written as a linear combination of the others, e.g.

$$\mathbf{v}_1 = k_2\mathbf{v}_2 + k_3\mathbf{v}_3 + \cdots + k_p\mathbf{v}_p\,. \qquad (11\text{-}9)$$

But this is equivalent to

$$\mathbf{v}_1 - k_2\mathbf{v}_2 - k_3\mathbf{v}_3 - \cdots - k_p\mathbf{v}_p = \mathbf{0}\,, \tag{11-10}$$

whereby the zero-vector is written as a linear combination of the vector set in which at least one of the coefficients are not 0, since $\mathbf{v}_1$ has the coefficient 1.

Conversely, assume that the zero-vector is written as a proper linear combination of the set of vectors, where one of the coefficients, for example the $\mathbf{v}_1$ coefficient k_1, is different from 0 (the same argument works for any of other coefficient). Then we have

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_p\mathbf{v}_p = \mathbf{0} \;\Leftrightarrow\; \mathbf{v}_1 = (-1)\frac{k_2}{k_1}\mathbf{v}_2 + \cdots + (-1)\frac{k_p}{k_1}\mathbf{v}_p\,. \tag{11-11}$$

Thus $\mathbf{v}_1$ is written as a linear combination of the other vectors and the proof is complete.

■

|||| Example 11.18 Linear Dependence

In the number space $\mathbb{R}^4$ the vectors $\mathbf{a} = (1,3,0,2)$, $\mathbf{b} = (-1,9,0,4)$ and $\mathbf{c} = (2,0,0,1)$ are given. Since

$$3\mathbf{a} - \mathbf{b} - 2\mathbf{c} = \mathbf{0}$$

the zero vector is written as a non-trivial linear combination of the three vectors. Thus they are linearly dependent.

11.4 Basis and Dimension of a Vector Space

A compelling argument for the introduction of a basis in a vector space is that all vectors in the vector space then can be written using coordinates. In a later section it is shown how problems of calculation can be simplified and standardized with vectors when we use coordinates. But in this section we will discuss the requirements that a basis should satisfy and investigate the consequences of these requirements.

A basis for a vector space consists of certain number of vectors, usually written in a definite order. A decisive task for the basis vectors is that they should span the vector space, but more precisely we want this task to be done with *as few vectors as possible*. In this case it turns out that all vectors in the vector space can be written *uniquely* as a linear combination of the basis vectors. And it is exactly the *coefficients* in the unique linear combination we will use as coordinates.

Let us start out from some characteristic properties about bases for geometric vectors in the plane.

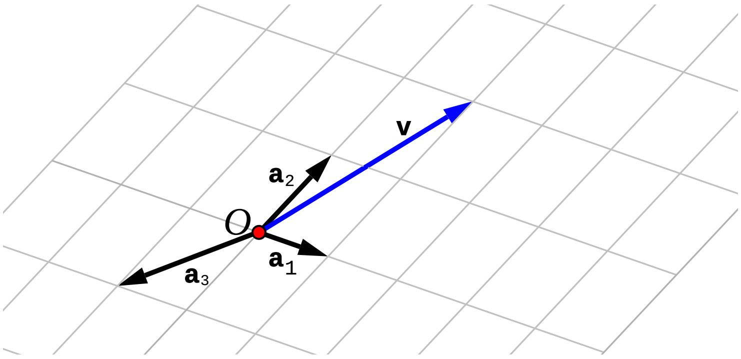

Figure 11.2: Coordinate system in the plane with the basis $(\mathbf{a}_1, \mathbf{a}_2)$

Consider the vector set $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ in Figure 11.2. There is no doubt that any other vector in the plane can be written as a linear combination of the three vectors. But the linear combination is not unique, for example the vector $\mathbf{v}$ can be written in these two ways:

$$\mathbf{v} = 2\mathbf{a}_1 + 3\mathbf{a}_2 - 1\mathbf{a}_3$$
$$\mathbf{v} = 1\mathbf{a}_1 + 2\mathbf{a}_2 + 0\mathbf{a}_3\,.$$

The problem is that the a-vectors are not linearly independent, for example $\mathbf{a}_3 = -\mathbf{a}_1 - \mathbf{a}_2$. But if we remove one of the vectors, e.g. $\mathbf{a}_3$, the set is linearly independent, and there is only one way of writing the linear combination

$$\mathbf{v} = 1\mathbf{a}_1 + 2\mathbf{a}_2\,.$$

We can summarize the characteristic properties of a basis for the geometric vectors in the plane thus:

1. any basis must consist of linearly independent vectors,
2. any basis must contain exactly two vectors (if there are more than two, they are linearly dependent, if there are less than two they do not span the plane), and
3. *every* set consisting of two linear independent vectors is a basis.

These properties can be transferred to other vector spaces. We embark on this now, and we start by the general definition of a basis.

|||| Definition 11.19 Basis

By a ***basis*** for a vector space V we understand a set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of vectors from V that satisfy:

1. $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ spans V.
2. $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is linearly independent.

When we discuss coordinates later, it will be necessary to consider the basis elements to have a define order, and so we will write them as an *ordered set*, denoted by using parentheses: $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$.

Here we should stop and check that the definition 11.19 does in fact satisfy our requirements of uniqueness of a basis. This is established in the following theorem.

|||| Theorem 11.20 Uniqueness Theorem

If a basis for a vector space V is given, any vector in V can then be written as a *unique* linear combination of the basis vectors.

|||| Proof

We give the idea in the proof by looking at a vector space V that has a basis consisting of three basis vectors $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ and assume that $\mathbf{v}$ is an arbitrary vector in V that in two ways can be written as a linear combination of the basis vectors. We can then write two equations

$$\begin{aligned} \mathbf{v} &= k_1\mathbf{a} + k_2\mathbf{b} + k_3\mathbf{c} \\ \mathbf{v} &= k_4\mathbf{a} + k_5\mathbf{b} + k_6\mathbf{c} \end{aligned} \tag{11-12}$$

By subtracting the lower equation from the upper equation in (11-12) we get the equation

$$\mathbf{0} = (k_1 - k_4)\mathbf{a} + (k_2 - k_5)\mathbf{b} + (k_3 - k_6)\mathbf{c}\,. \tag{11-13}$$

Since $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are linearly independent, the zero vector can only be written as an improper linear combination of these, therefore all coefficients in (11-12) are equal to 0, yielding $k_1 = k_4$, $k_2 = k_5$ and $k_3 = k_6$. But then the two ways $\mathbf{v}$ has been written as linear combinations of the basis vectors, is in reality the same, there is only one way!

This reasoning is immediately extendable to a basis consisting of an arbitrary number of basis vectors.

■

We now return to the fact that every basis for geometric vectors in the plane contains *two* linearly independent basis vectors, and that similarly for geometric vectors in space the basis must consist of *three* linearly independent basis vectors. It turns out that the fixed number of basis vectors is a property of all vector spaces with a basis, and this makes it possible to talk about the *dimension* of a vector space that has a basis. To prove the property we need a lemma.

|||| Lemma 11.21

If a vector space V has a basis consisting of n basis vectors then every set from V that contains more than n vectors will be linearly dependent.

|||| Proof

To get a grasp of the proof's underlying idea, consider a vector space V that has a basis consisting of two vectors $(\mathbf{a}, \mathbf{b})$, and investigate three arbitrary vectors $\mathbf{c}, \mathbf{d}$ and $\mathbf{e}$ from V. We prove that the three vectors necessarily must be linearly independent.

Since $(\mathbf{a}, \mathbf{b})$ is a basis for V, we can write three equations

$$\begin{aligned} \mathbf{c} &= c_1\mathbf{a} + c_2\mathbf{b} \\ \mathbf{d} &= d_1\mathbf{a} + d_2\mathbf{b} \\ \mathbf{e} &= e_1\mathbf{a} + e_2\mathbf{b} \end{aligned} \tag{11-14}$$

Consider further the zero vector written as the following linear combination

$$x_1\mathbf{c} + x_2\mathbf{d} + x_3\mathbf{e} = \mathbf{0}\,, \tag{11-15}$$

which by substitution of the equations (11-14) into (11-15) is equivalent to

$$(x_1c_1 + x_2d_1 + x_3e_1)\mathbf{a} + (x_1c_2 + x_2d_2 + x_3e_2)\mathbf{b} = \mathbf{0}\,. \tag{11-16}$$

Since the zero vector only can be obtained as a linear combination of $\mathbf{a}$ and $\mathbf{b}$, if every coefficient is equal to 0, (11-16) is equivalent to the following system of equations

$$\begin{aligned} c_1x_1 + d_1x_2 + e_1x_3 &= 0 \\ c_2x_1 + d_2x_2 + e_2x_3 &= 0 \end{aligned} \tag{11-17}$$

This is a homogeneous system of linear equations in which the number of equations is less than the number of unknowns. Therefore the system of equations has infinitely many solutions, which means that (11-16) not only is obtainable with $x_1 = 0, x_2 = 0$ and $x_3 = 0$. Thus we have shown that the ordered set $(\mathbf{c}, \mathbf{d}, \mathbf{e})$ is linearly dependent.

In general: Assume that the basis V consists of n vectors, and that m vectors from V where $m > n$ are given. By following the same procedure as above a homogeneous system of linear equations emerges with n equations in m unknowns that, because $m > n$, similarly has infinitely many solutions. By this it is shown that the m vectors are linearly dependent.

■

Then we are ready to give the following important theorem:

|||| Theorem 11.22 The Number of Basis Vectors

If a vector space V has a basis consisting of n basis vectors, then every basis for V will consist of n basis vectors.

|||| Proof

Assume that V has two bases with different numbers of big(asis vectors. We denote the basis with the least number of basis vectors a and the one with largest number b. According to Lemma 11.21 the b-basis vectors must be linearly dependent, and this contradicts that they form a basis. The assumption that V can have two bases with different numbers of basis vectors, must therefore be untrue.

■

That the number of basis vectors according to theorem 11.22 is a *property* of vector spaces with a basis, motivates the introduction of the concept of dimension:

|||| Definition 11.23 Dimension

By the dimension of a vector space V that has a basis b, we understand the number of vectors in b. If this number is n, one says that V is n-dimesional and write

$$\dim(V) = n\,. \tag{11-18}$$

Remark: There are vector spaces that do not have a finite basis, see Section 11.7.2 below.

|||| Example 11.24 Dimension of Geometric Vector Spaces

Luckily the definition 11.23 confirms an intuitive feeling that the set of geometric vectors in the plane has the dimension two and that the set of geometric vectors in space has the dimension three!

|||| Example 11.25 The Standard Basis for Number Spaces

An arbitrary vector $\mathbf{v} = (a, b, c, d)$ in $\mathbb{R}^4$ or in $\mathbb{C}^4$ (that is in $\mathbb{L}^4$) can in an obvious way be written as a linear combination of four particular vectors in $\mathbb{L}^4$

$$\mathbf{v} = a\,(1, 0, 0, 0) + b\,(0, 1, 0, 0) + c\,(0, 0, 1, 0) + d\,(0, 0, 0, 1)\,. \tag{11-19}$$

We put $\mathbf{e}_1 = (1, 0, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0, 0)$, $\mathbf{e}_3 = (0, 0, 1, 0)$ and $\mathbf{e}_4 = (0, 0, 0, 1)$ and conclude using (11-19) that the ordered set $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ spans $\mathbb{L}^4$.

Since we can see that none of the vectors can be written as a linear combination of the others, the set is linearly independent, and $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ is thereby a basis for $\mathbb{L}^4$. This particular basis is called ***standard basis*** for $\mathbb{L}^4$. Since the number of basis vectors in the standard e-basis is four, $\dim(\mathbb{L}^4) = 4$.

This can immediately be generalized to $\mathbb{L}^n$: For every n the set $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ where

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \mathbf{e}_n = (0, 0, 0, \ldots, 1)$$

is a basis for $\mathbb{L}^n$. This is called ***standard basis*** for $\mathbb{L}^n$. It is noticed that $\dim(\mathbb{L}^n) = n$.

|||| Example 11.26 Standard Basis for Matrix Spaces

By ***standard basis*** for the vector space $\mathbb{R}^{2\times 3}$ or $\mathbb{C}^{2\times 3}$, we understand the matrix set

$$\left(\begin{bmatrix}1&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&1&0\\0&0&0\end{bmatrix},\ldots,\begin{bmatrix}0&0&0\\0&0&1\end{bmatrix}\right) \tag{11-20}$$

Similarly we define a *standard basis* for an arbitrary matrix space $\mathbb{R}^{m\times n}$ and for an arbitrary matrix space $\mathbb{C}^{m\times n}$.

|||| Exercise 11.27

Explain that the matrix set, which in Example 11.26 is referred to as the standard basis for $\mathbb{R}^{2\times 3}$, is in fact a *basis* for this vector space.

|||| Example 11.28 The Monomial Basis for Polynomial Spaces

In the vector space $P_2(\mathbb{R})$ of real polynomials of at most 2nd degree, the ordered set $(1, x, x^2)$ is a basis. This is demonstrated in the following way.

1. Every polynomial $P(x) \in P_2(\mathbb{R})$ can be written in the form

$$P(x) = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 \,,$$

   that is as a linear combination of the three vectors in the set.

2. The set $\{1, x, x^2\}$ is linearly independent, since the equation

$$a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 = 0 \text{ for every } x$$

   according to the ***identity theorem for polynomials*** is only satisfied if all the coefficients a_0, a_1 and a_2 are equal to 0.

A *monomial* is a polynomial with only one term. Hence, the ordered set $(1, x, x^2)$ is called the ***monomial basis*** for $P_2(\mathbb{R})$, and $\dim(P_2(\mathbb{R})) = 3$.

For every n the ordered set $(1, x, x^2, \ldots, x^n)$ is a basis for $P_n(\mathbb{R})$, and is called the *monomial basis* for $P_n(\mathbb{R})$. Therefore we have that $\dim(P_n(\mathbb{R})) = n + 1$.

Similarly the ordered set $(1, z, z^2, \ldots, z^n)$ is a basis for $P_n(\mathbb{C})$, it is called *monomial basis* for $P_n(\mathbb{C})$. Therefore we have that $\dim(P_n(\mathbb{C})) = n + 1$.

In the set of plane geometric vectors one can choose *any* pair of two linearly independent vectors as basis. Similarly in 3-space *any* set of three linear independent vectors is a basis. We end the section by transferring this to general n-dimensional vector spaces:

|||| Theorem 11.29 Sufficient Conditions for a Basis

In an n-dimensional vector space V, an arbitrary set of n linearly independent vectors from V constitutes a basis for V .

|||| Proof

Since V is assumed to be n-dimensional, it must have a basis b consisting of n basis vectors. Let the a-set $(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n)$ be an arbitrary linearly independent set of vectors from V. The set is then a basis for V if it spans V. Suppose this is not the case, and let $\mathbf{v}$ be a vector V that does not belong to $\text{span}\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n\}$. Then $(\mathbf{v}, \mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n)$ must be linearly independent, but this contradicts theorem 11.21 since there are $n+1$ vectors in the set. Therefore the assumption that the a-set does not span V must be untrue, and it must accordingly be a basis for V.

■

|||| Exercise 11.30

Two geometric vectors $\mathbf{a} = (1, -2, 1)$ and $\mathbf{b} = (2, -2, 0)$ in 3-space are given. Determine a vector $\mathbf{c}$ such that the ordered set $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is a basis for the set of space vectors.

|||| Exercise 11.31

In the 4-dimensional vector space $\mathbb{R}^{2\times 2}$, consider the vectors

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ of } \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{11-21}$$

Explain why the ordered set $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ is a linearly independent set, and complement the set with a 2×2 matrix $\mathbf{D}$ such that $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ is a basis for $\mathbb{R}^{2\times 2}$.

11.5 Vector Calculations Using Coordinates

Coordinates are closely connected to the concept of a basis. When a basis is chosen for a vector space, any vector in the vector space can be described with the help of its co-ordinates with respect to the chosen basis. By this we get a particularly practical alternative to the calculation operations, addition and multiplication by a scalar, which originally are defined from the 'anatomy' of the specific vector space. Instead of carrying out these particularly defined operations we can implement number calculations with the coordinates that correspond to the chosen basis. In addition it turns out that we can simplify and standardize the solution of typical problems that are common to all vector spaces. But first we give a formal introduction of coordinates with respect to a chosen basis.

|||| Definition 11.32 Coordinates with Respect to a Given Basis

In an n-dimensional vector space V the basis $a = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ and a vector $\mathbf{x}$ are given. We consider the unique linear combination of the basis vectors that according to 11.20 is a way of writing $\mathbf{x}$:

$$\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n\,. \tag{11-22}$$

The coefficients $x_1, x_2, \ldots, x_n$ in (11-22) are denoted $\mathbf{x}$'s *coordinates with respect to the basis* a, or $\mathbf{x}$'s a-coordinates, and they are gathered in a *coordinate vector* as follows:

$${}_a\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{11-23}$$

|||| Example 11.33 Coordinates with Respect to a New Basis

In the number space $\mathbb{R}^3$ a basis a is given by $((0,0,1),(1,2,0),(1,-1,1))$. Furthermore the vector $\mathbf{v} = (7,2,6)$ is given. Since

$$2 \cdot (0,0,1) + 3 \cdot (1,2,0) + 4 \cdot (1,-1,1) = (7,2,6)$$

we see that

$$_{\mathrm{a}}\mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}.$$

The vector $(7,2,6)$ therefore has the a-coordinates $(2,3,4)$.

In order to be able to manipulate the coordinates of several vectors in various arithmetic operations we will need the following important theorem.

|||| Theorem 11.34 The Coordinate Theorem

In a vector space V two vectors $\mathbf{u}$ and $\mathbf{v}$ plus a real number k are given. In addition an arbitrary basis a is chosen. The two arithmetic operations $\mathbf{u}+\mathbf{v}$ of $k\,\mathbf{u}$ can then be carried out using the a-coordinates like this:

1. $_{\mathrm{a}}(\mathbf{u}+\mathbf{v}) = {_{\mathrm{a}}\mathbf{u}} + {_{\mathrm{a}}\mathbf{v}}$
2. $_{\mathrm{a}}(k\mathbf{u}) = k\,{_{\mathrm{a}}\mathbf{u}}$

In other words: The coordinates for a vector sum are obtained by adding the coordinates for the vectors, and the coordinates for a vector multiplied by a number are the coordinates of the vector multiplied by the number.

|||| Proof

See the proof for the corresponding theorem for geometric vectors in 3-space, Theorem 10.38. The proof for the general case is obtained as a simple extension.

■

|||| Example 11.35 Vector Calculation Using Coordinates

We now carry out a vector calculation using coordinates. The example is not particularly mathematically interesting, but we carry it out in detail in order to demonstrate the technique of Theorem 11.34.

There are given three polynomials in the vector space $P_2(\mathbb{R})$:

$$R(x) = 2 - 3x - x^2\,, S(x) = 1 - x + 3x^2 \text{ and } T(x) = x + 2x^2\,.$$

The task is now to determine the polynomial $P(x) = 2R(x) - S(x) + 3T(x)$. We choose to carry this out using coordinates for the polynomials with respect to the monomial basis for $P_2(\mathbb{R})$.

$$\begin{aligned} {}_{\mathrm{m}}P(x) &= {}_{\mathrm{m}}\big(2R(x) - S(x) + 3T(x)\big) \\ &= {}_{\mathrm{m}}(2R(x)) + {}_{\mathrm{m}}(-S(x)) + {}_{\mathrm{m}}(3T(x)) \\ &= 2\,{}_{\mathrm{m}}R(x) - {}_{\mathrm{m}}S(x) + 3\,{}_{\mathrm{m}}T(x) \\ &= 2\begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}. \end{aligned}$$

We translate the resulting coordinate vector to the wanted polynomial:

$$P(x) = 3 - 2x + x^2\,.$$

11.6 On the Use of Coordinate Matrices

When we embark on problems with vectors and use their coordinates with respect to a given basis it often leads to a system of linear equations which we then solve by matrix calculations. One matrix is of particular importance, viz. the matrix that is formed by gathering the coordinate columns of more vectors in a *coordinate matrix*:

|||| Explanation 11.36 Coordinate Matrix for a Vector Set

If in a n-dimensional vector space V a basis a exists, and a set of m numbered vectors is given, then the ***a*-coordinate matrix** is formed by gathering the a-coordinate columns in the given order to form an $m\times n$ matrix.

By way of example consider a set of three vectors in $\mathbb{R}^2$: $((1,2),(3,4),(5,6))$. The coordinate matrix of the set with respect to the standard e-basis for $\mathbb{R}^2$ is the 2×3-matrix

$$\begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

We will now show how coordinate matrices emerge in series of examples which we, for the sake of variation, take from different vector spaces. The methods can directly be used on other types of vector spaces, and after each example the method is demonstrated in a concentrated and general form.

It is important for your own understanding of the theory of vector spaces that you practice and realize how coordinate matrices emerge in reality when you start on typical problems.

11.6.1 Whether a Vector is a Linear Combination of Other Vectors

In $\mathbb{R}^4$ we are given four vectors

$$\begin{aligned} \mathbf{a}_1 &= (1,1,1,1) \\ \mathbf{a}_2 &= (1,0,0,1) \\ \mathbf{a}_3 &= (2,3,1,4) \\ \mathbf{b} &= (2,-2,0,1) \end{aligned}$$

Problem: Investigate if $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$.

Solution: We will investigate whether we can find $x_1, x_2, x_3 \in \mathbb{R}$ such that

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b}\,. \tag{11-24}$$

By theorem 11.34 we can rewrite (11-24) as the e-coordinate vector equation

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 3 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 0 \\ 1 \end{bmatrix}$$

which is equivalent to the system of linear equations

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 2 \\ x_1 + 3x_3 &= -2 \\ x_1 + x_3 &= 0 \\ x_1 + x_2 + 4x_3 &= 1 \end{aligned}$$

We form the augmented matrix of the system of equations and give (without further details) its reduced row echelon form

$$\mathbf{T} = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 0 & 3 & -2 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 4 & 1 \end{bmatrix} \Rightarrow \operatorname{rref}(\mathbf{T}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{11-25}$$

From (11-25) it is seen that the rank of the coefficient matrix of the system of equations is 3, while the rank of the augmented matrix is 4. The system of equations has therefore no solutions. This means that (11-24) cannot be solved. We conclude

$$\mathbf{b} \notin \operatorname{span}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}.$$

|||| Method 11.37 Linear Combination

You can decide whether a given vector $\mathbf{b}$ is a linear combination of other vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ by solving the system of linear equations which has the augmented matrix that is equal to the coordinate matrix for $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p, \mathbf{b})$ with respect to a given basis.

NB: In general there can be none, one or infinitely many ways a vector can be written as linear combinations of the others.

11.6.2 Whether Vectors are Linearly Dependent

We consider in the vector space $\mathbb{R}^{2\times 3}$ the three matrices:

$$\mathbf{A}=\begin{bmatrix}1&0&3\\0&2&2\end{bmatrix},\ \mathbf{B}=\begin{bmatrix}2&1&0\\0&3&1\end{bmatrix}\text{ and }\mathbf{C}=\begin{bmatrix}-1&-2&9\\0&0&4\end{bmatrix}. \tag{11-26}$$

Problem: Investigate whether the three matrices are linearly dependent.

Solution: We use theorem 11.17 and try to find three real numbers x_1 , x_2 of x_3 that are not all equal to 0, but which satisfy

$$x_1\mathbf{A}+x_2\mathbf{B}+x_3\mathbf{C}=\begin{bmatrix}0&0&0\\0&0&0\end{bmatrix}. \tag{11-27}$$

By theorem 11.34 we can rewrite (11-27) as the e-coordinate vector equation

$$x_1\begin{bmatrix}1\\0\\3\\0\\2\\2\end{bmatrix}+x_2\begin{bmatrix}2\\1\\0\\0\\3\\1\end{bmatrix}+x_3\begin{bmatrix}-1\\-2\\9\\0\\0\\4\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix}$$

That is equivalent to the homogeneous system of linear equations with the augmented matrix that here is written together with reduced row echelon form (details are omitted):

$$\mathbf{T}=\begin{bmatrix}1&2&-1&0\\0&1&-2&0\\3&0&9&0\\0&0&0&0\\2&3&0&0\\2&1&4&0\end{bmatrix}\Rightarrow \operatorname{rref}(\mathbf{T})=\begin{bmatrix}1&0&3&0\\0&1&-2&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}. \tag{11-28}$$

From (11-28) we see that both the coefficient matrix and the augmented matrix have the rank 2, and since the number of unknowns is larger, viz. 3, we conclude that Equation (11-27) has infinitely many solutions , see Theorem 6.33. Hence the three matrices are linearly dependent. For instance, from $\operatorname{rref}(\mathbf{T})$ one can derive that

$$-3\mathbf{A}+2\mathbf{B}+\mathbf{C}=\begin{bmatrix}0&0&0\\0&0&0\end{bmatrix}.$$

|||| **Method 11.38 Linear Dependence or Independence**

One can decide whether the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ are linearly dependent by solving the linear homogenous system of linear equations with the augmented matrix that is equal to the coordinate matrix for $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p, \mathbf{0})$ with respect to a given basis.

NB: Since the system of equations is homogeneous, there will be either one solution or infinitely many solutions. If the rank of the coordinate matrix is equal to p, there is one solution, and this solution must be the zero solution, and the p vectors are therefore linearly independent. If the rank of the coordinate matrix is less than p, there are infinitely many solutions, including non-zero solutions, and the p vectors are therefore linearly dependent.

11.6.3 Whether a Set of Vectors is a Basis

In an n-dimensional vector space we require n basis vectors, see theorem 11.22. When one has asked whether a given set of vectors can be a basis, one can immediately conclude that this is not the case if the number of vectors in the set is not equal to n. But if there *are* n vectors in the set according to theorem 11.29 we need only investigate whether the set is linear independent, and for this we already have method 11.38. However we can in an interesting way develop the method further by using the determinant of the coordinate matrix of the vector set!

Let us e.g. investigate whether the polynomials

$$P_1(x) = 1 + 2x^2,\; P_2(x) = 2 - x + x^2 \;\text{ of }\; P_3(x) = 2x + x^2$$

form a basis for $P_2(\mathbb{R})$. Since $\dim(P_2(\mathbb{R}))$ = 3, the number of polynomials is compatible with being a basis. In order to investigate whether they also are linearly independent, we use their coordinate vectors with respect to the ***monomial basis*** and consider the equation:

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The vectors are linearly independent if and only if the only solution is the trivial solution $x_1 = x_2 = x_3 = 0$. The equation is equivalent to a homogeneous system of linear equations consisting of 3 equations in 3 unknowns. The coefficient matrix and the augmented matrix of the system are:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 2 \\ 2 & 1 & 1 \end{bmatrix} \text{ and } \mathbf{T} = \left[\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 2 & 1 & 1 & 0 \end{array}\right].$$

As for every homogeneous system of linear equations the right hand side of the augmented matrix consists of only 0's, therefore $\rho(\mathbf{A}) = \rho(\mathbf{T})$, and thus solutions *do* exist. There is one solution exactly when $\rho(\mathbf{A})$ is equal to the number of unknowns, that is 3. And this solution must be the zero solution $x_1 = x_2 = x_3 = 0$, since L_{hom} always contains the zero solution.

Here we can use that $\mathbf{A}$ is a square matrix and thus has a *determinant*. $\mathbf{A}$ has full rank exactly when it is *invertible*, that is when $\det(\mathbf{A}) \neq 0$.

Since a calculation shows that $\det(\mathbf{A}) = 5$ we conclude that $\big(P_1(x), P_2(x), P_3(x)\big)$ constitutes a basis for $P_2(\mathbb{R})$.

|||| Method 11.39 Proof of a Basis, given n vectors

Given an n-dimensional vector space V. To determine whether a vector set consisting of n vectors $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ is a basis for V, we only need to investigate whether the set is linearly independent. A particular option for this investigation occurs because the coordinate matrix of the vector set is a square $n \times n$ matrix:

The set constitutes a basis for V exactly when the determinant of the coordinate matrix of the set with respect to a basis a is non-zero, in short

$$(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) \text{ is a basis} \Leftrightarrow \det\big(\big[\,{}_a\mathbf{v}_1 \;\; {}_a\mathbf{v}_2 \;\; \cdots \;\; {}_a\mathbf{v}_n\,\big]\big) \neq 0\,. \tag{11-29}$$

11.6.4 To Find New Coordinates when the Basis is Changed

An important technical problem for the advanced use of linear algebra is to be able to calculate new coordinates for a vector when a new basis is chosen. In this context a particular *change of basis matrix* plays an important role. We now demonstrate how basis matrices emerge.

In a 3-dimensional vector space V a basis a is given. We now choose a new basis b that is determined by the a-coordinates of the basis vectors:

$$_{\mathrm{a}}\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},\ _{\mathrm{a}}\mathbf{b}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \text{ and } _{\mathrm{a}}\mathbf{b}_3 = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}.$$

Problem 1: Determine the a-coordinates for a vector $\mathbf{v}$ given by the b-coordinates as:

$$_{\mathrm{b}}\mathbf{v} = \begin{bmatrix} 5 \\ -4 \\ -1 \end{bmatrix}. \tag{11-30}$$

Solution: The expression (11-30) corresponds to the vector equation

$$\mathbf{v} = 5\mathbf{b}_1 - 4\mathbf{b}_2 - 1\mathbf{b}_3$$

which we below first convert to an a-coordinate vector equation, re-writing the right hand side as a matrix-vector product, before finally computing the result:

$$\begin{aligned} _{\mathrm{a}}\mathbf{v} &= 5\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 4\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - 1\begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \end{bmatrix}\begin{bmatrix} 5 \\ -4 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix}. \end{aligned}$$

Notice that the 3×3-matrix in the last equation is the coordinate matrix for the b-basis vectors with respect to basis a. It plays an important role, since we apparently can determine the a-coordinates for $\mathbf{v}$ by multiplying b-coordinate vector for $\mathbf{v}$ on the left by this matrix! Therefore the matrix is given the name *change of basis matrix*. The property of this matrix is that it translates b-coordinates to a-coordinates, and it is given the symbol $_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}$. The coordinate change relation can then be written in this convenient way

$$_{\mathrm{a}}\mathbf{v} = {}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}\,{}_{\mathrm{b}}\mathbf{v}\,. \tag{11-31}$$

Problem 2: Determine the b-coordinates for a vector $\mathbf{u}$ that has a-coordinates:

$$_{\mathrm{a}}\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}. \tag{11-32}$$

Solution: Since ${}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}$ is the coordinate matrix for a basis, it is *invertible*, and thus has an inverse matrix. We therefore use the coordinate change relation (11-31) as follows:

$$\begin{aligned}
{}_{\mathrm{a}}\mathbf{u} &= {}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}\,{}_{\mathrm{b}}\mathbf{u} \;\Leftrightarrow \\
{}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}{}^{-1}\,{}_{\mathrm{a}}\mathbf{u} &= {}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}{}^{-1}\,{}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}\,{}_{\mathrm{b}}\mathbf{u} \;\Leftrightarrow \\
{}_{\mathrm{b}}\mathbf{u} &= {}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}{}^{-1}\,{}_{\mathrm{a}}\mathbf{u} \;\Leftrightarrow \\
{}_{\mathrm{b}}\mathbf{u} &= \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ -4 \\ -3 \end{bmatrix}.
\end{aligned}$$

|||| Method 11.40 Coordinate Change when the Basis is Changed

When a basis a is given for a vector space, and when a new basis b is known by the a-coordinates of its basis vectors, the ***change of basis matrix*** ${}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}$ is identical to the a-coordinate matrix for b-basis vectors.

1. If b-coordinates for a vector $\mathbf{v}$ are known, these a-coordinates can be found by the matrix-vector product:
$$ {}_{\mathrm{a}}\mathbf{v} = {}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}\,{}_{\mathrm{b}}\mathbf{v}\,. $$

2. Conversely, if the a-coordinates for $\mathbf{v}$ are known, the b-coordinates can be found by the matrix-vector product:
$$ {}_{\mathrm{b}}\mathbf{v} = {}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}{}^{-1}\,{}_{\mathrm{a}}\mathbf{v}\,. $$

   In short the change of basis matrix that translates a-coordinates to b-coordinates, is the inverse of the change of basis matrix that translate b-coordinates to a-coordinates:
$$ {}_{\mathrm{b}}\mathbf{M}_{\mathrm{a}} = \left({}_{\mathrm{a}}\mathbf{M}_{\mathrm{b}}\right)^{-1}. $$

11.7 Subspaces

Often you encounter that a subset of a vector space is itself a vector space. In Figure 11.3 are depicted two position vectors $\overrightarrow{OP}$ and $\overrightarrow{OQ}$ that span the plane F:

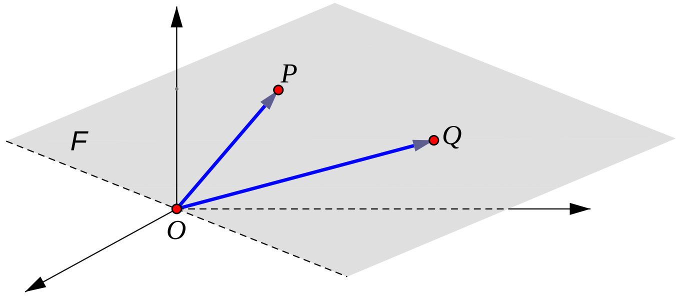

Figure 11.3: A plane through the origin interpreted as a *subspace* in space

Since span$\{\overrightarrow{OP}, \overrightarrow{OQ}\}$ can be considered to be a (2-dimensional) vector space in its own right, it is named a *subspace* of the (3-dimensional) vector space of position vectors in space.

|||| Definition 11.41 Subspace

A subset U of a vector space V is called a ***subspace*** of V if U is itself a vector space.

In any vector space V one can immediately point to two subspaces:
1) V is in itself a subspace of V.
2) The set $\{\mathbf{0}\}$ is a subspace of V.
These subspaces are called the *trivial* subspaces in V.

When one must check whether a subset is a subspace, one only has to check whether the stability requirements are satisfied:

|||| Theorem 11.42 Sufficient Conditions for a Subspace

A non-empty subset U of a vector space V is a subspace of V if U is *stable* with respect to addition and multiplication by a scalar. This means

1. The sum of two vectors from U belongs to U.
2. The product of a vector in U with a scalar belongs to U.

|||| Proof

Since U satisfies the two stability requirements in 11.1, it only remains to show that U also satisfies the eight arithmetic rules in the definition. But this is evident since all vectors in U are also vectors in V where the rules apply.

■

|||| Example 11.43 Basis for a Subspace

We consider a subset M_1 of $\mathbb{R}^{2\times 2}$, consisting of matrices of the type

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix} \tag{11-33}$$

where a and b are arbitrary real numbers. We try to add two matrices of the type (11-33)

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 4 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 4 \end{bmatrix}$$

and we multiply one of type (11-33) by a scalar

$$-3 \begin{bmatrix} 2 & -3 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} -6 & 9 \\ 9 & -6 \end{bmatrix}.$$

in both cases the resulting matrix is of type (11-33) and it is obvious that this would also apply had we used other examples. Therefore M_1 satisfies the stability requirements for a vector space. Thus it follows from theorem 11.42 that M_1 is a subspace of $\mathbb{R}^{2\times 2}$.

Further remark that M_1 is spanned by two linear independent 2×2 matrices since

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Therefore M_1 is a 2-dimensional subspace of $\mathbb{R}^{2\times 2}$, and a possible basis for M_1 is given by

$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right).$$

|||| Example 11.44 A Subset which is Not a Subspace

The subset M_2 of $\mathbb{R}^{2\times 2}$ consists of all matrices of the type

$$\begin{bmatrix} a & b \\ a\cdot b & 0 \end{bmatrix} \tag{11-34}$$

where a and b are arbitrary real numbers. We try to add two matrices of the type (11-34)

$$\begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ 6 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 8 & 0 \end{bmatrix}.$$

Since $8 \neq 3\cdot 5$, this matrix is not of the type (11-34). Therefore M_2 is not stable under linear combinations, and cannot be a subspace.

11.7.1 About Spannings as Subspaces

|||| Theorem 11.45 Spannnings of Subspaces

For arbitrary vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ in vector space V, the set $\operatorname{span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\}$ is a subspace of V.

|||| Proof

The stability requirements are satisfied because 1) the sum of two linear combinations of the p vectors in itself is a linear combination of them and 2) a linear combination of the p vectors multiplied by a scalar in itself is a linear combination of them. The rest follows from Theorem 11.42.

■

The solution set L_{hom} for a homogeneous system of linear equations with n unknowns is always a subspace of the number space $\mathbb{R}^n$ and the dimension of the subspace is the same as the number of free parameters in L_{hom}. We show an example of this below.

|||| Example 11.46 L_{hom} is a Subspace

The following homogeneous system of linear equations of 3 equations in 5 unknowns

$$\begin{aligned} x_1 + 2\,x_3 - 11\,x_5 &= 0 \\ x_2 + 4\,x_5 &= 0 \\ x_4 + x_5 &= 0 \end{aligned}$$

has the solution set (details are omitted):

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = t_1 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} 11 \\ -4 \\ 0 \\ -1 \\ 1 \end{bmatrix} \quad \text{where } t_1, t_2 \in \mathbb{R}\,. \tag{11-35}$$

We see that L_{hom} is a span of two vectors in $\mathbb{R}^5$. Then it is according to theorem 11.45 a subspace of $\mathbb{R}^5$. Since the two vectors evidently are linearly independent, L_{hom} is a 2-dimensional subspace of $\mathbb{R}^5$, with a basis

$$\big(\,(-2, 0, 1, 0, 0)\,,(11, -4, 0, -1, 1)\,\big)\,.$$

In the following example we will establish a method for how one can determine a basis for a subspace that is spanned by a number of given vectors in a subspace.

Consider in $\mathbf{R}^3$ four vectors

$$\mathbf{v}_1 = (1, 2, 1),\ \mathbf{v}_2 = (3, 0, -1),\ \mathbf{v}_3 = (-1, 4, 3) \text{ and } \mathbf{v}_4 = (8, -2, -4)\,.$$

We wish to find a basis for the subspace

$$U = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}\ .$$

Let $\mathbf{b} = (b_1, b_2, b_3)$ be an arbitrary vector in U. We thus assume that the following vector equation has a solution:

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + x_4\mathbf{v}_4 = \mathbf{b}\,. \tag{11-36}$$

By substitution of the five vectors into (11-36), it is seen that (11-36) is equivalent to an inhomogeneous system of linear equations with the augmented matrix:

$$\mathbf{T} = \left[\begin{array}{rrrr|r} 1 & 3 & -1 & 8 & b_1 \\ 2 & 0 & 4 & -2 & b_2 \\ 1 & -1 & 3 & -4 & b_3 \end{array}\right] \Rightarrow \operatorname{rref}(\mathbf{T}) = \left[\begin{array}{rrrr|r} 1 & 0 & 2 & -1 & c_1 \\ 0 & 1 & -1 & 3 & c_2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{11-37}$$

Here c_1 is placeholder for the number that b_1 has been transformed into following the row operations leading to the reduced row echelon form $\operatorname{rref}(\mathbf{T})$. Similarly for c_2. Remark that b_3 after the row operations must be transformed into 0, or else (x_1, x_2, x_3, x_4) could not be a solution as we have assumed.

But it is in particular the leading 1's in $\operatorname{rref}(\mathbf{T})$ on which we focus! They show that $\mathbf{v}_1$ and $\mathbf{v}_2$ span all of U, and that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linear independent. We can convince ourselves of both by considering equation (11-36) again.

First: Suppose we had only asked whether $\mathbf{v}_1$ and $\mathbf{v}_2$ span all of U. Then we should have omitted the terms with $\mathbf{v}_3$ and $\mathbf{v}_4$ from (11-36), and then we would have obtained:

$$\operatorname{rref}(\mathbf{T}_2) = \left[\begin{array}{rr|r} 1 & 0 & c_1 \\ 0 & 1 & c_2 \\ 0 & 0 & 0 \end{array}\right]$$

that shows that $c_1\,\mathbf{v}_1 + c_2\,\mathbf{v}_2 = \mathbf{b}$, and that $\mathbf{v}_1$ and $\mathbf{v}_2$ then span all of U.

Secondly: Suppose we had asked whether $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. Then we should have omitted the terms with $\mathbf{v}_3$ and $\mathbf{v}_4$ from (11-36), and put $\mathbf{b} = \mathbf{0}$. And then we would have got:

$$\operatorname{rref}(\mathbf{T}_3) = \left[\begin{array}{rr|r} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

That shows that the zero vector can only be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ if both of the coefficients x_1 and x_2 are 0. And thus we show that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. In total we have shown that $(\mathbf{v}_1, \mathbf{v}_2)$ is a basis for U.

The conclusion is that a basis for U can be singled out by the leading 1's in $\operatorname{rref}(\mathbf{T})$, see (11-37). The right hand side in $\operatorname{rref}(\mathbf{T})$ was meant to serve our argument but its contribution is now unnecessary. Therefore we can summarize the result as the following method:

|||| **Method 11.47 About refining a Spanning Set to a Basis**

When, in a vector space V, for which a basis a has been chosen, one wishes to find a basis for the subspace

$$U = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$$

everything can be read from

$$\text{rref}\left(\begin{bmatrix} {}_a\mathbf{v}_1 & {}_a\mathbf{v}_2 & \ldots & {}_a\mathbf{v}_p \end{bmatrix}\right). \tag{11-38}$$

If in the i'th column in (11-38) there are no leading 1's, then $\mathbf{v}_i$ is deleted from the set $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$. The set reduced in this way is a basis for U.

Since the number of leading 1's in (11-38) is equal to the number of basis vectors in the chosen basis for U, it follows that

$$\text{Dim}(U) = \rho\left(\begin{bmatrix} {}_a\mathbf{v}_1 & {}_a\mathbf{v}_2 & \ldots & {}_a\mathbf{v}_p \end{bmatrix}\right). \tag{11-39}$$

11.7.2 Infinite-Dimensional Vector Space

Before we end this eNote, that has cultivated the use of bases and coordinates, we must admit that not all vector spaces have a basis. Viz. there exist ***infinite-dimensional vector spaces***.

This we can see through the following example:

|||| **Example 11.48 Infinite-Dimensional Vector Spaces**

All polynomials in the vector space $P_n(\mathbb{R})$ are continuous functions, therefore $P_n(\mathbb{R})$ is an n+1 dimensional subspace of the vector space $C^0(\mathbb{R})$ of all real continuous functions. Now consider $P(\mathbb{R})$, the set of all real polynomials, that for the same reason is also a subspace of $C^0(\mathbb{R})$. But $P(\mathbb{R})$ must be *infinite-dimensional*, since it has $P_n(\mathbb{R})$, for every n, as a subspace. For the same reason $C^0(\mathbb{R})$ must also be infinite-dimensional.

|||| Exercise 11.49

By $C^1(\mathbb{R})$ is understood the set of all differentiable functions, with $\mathbb{R}$ as their domain, and with continuous derivatives in $\mathbb{R}$.

Explain why $C^1(\mathbb{R})$ is an infinite-dimensional subspace of $C^0(\mathbb{R})$.