7 Vector spaces

7.1 Sets and fields

A vector space is, loosely speaking, a set of objects that can be multiplied by scalars and added. We will see a lot of examples later, but at this point I would like first to talk a little about *sets* and to clarify why I choose to call some constants in this course "scalars" instead of more familiar "numbers." Anyway, every time when I talked about scalars I meant some real numbers.

The most basic structure in mathematics is a *set*. Since it is *most* basic, we cannot define it[1] in a proper mathematical sense, and I describe a set as a collections of some objects whose nature is not important. Usually sets are denoted using the curvy brackets:

$$A = \{a, b, c, d\},$$

which means that set A consists of 4 distinct elements a, b, c, d, note that the order of these elements is not important. While working with sets, I use the notation $a \in A$ to state that element a belongs to set A and $e \notin A$ for element e does not belong to A. A very special set is the *empty set* $\varnothing = \{\,\}$ which contains no elements. A subset B of a set A is defined to be a set such that each element of B is also an element of A, the notation is $B \subseteq A$. For example, if $B = \{a, b, c\}$ then B is a subset of the defined above set A, whereas if $B = \{a, b, e\}$ then B is not a subset of A. By convention empty set is a subset of any set: $\varnothing \subseteq A$. It is always true that $A \subseteq A$ for any set A. A set B is called a *proper subset of* A, if $B \subseteq A$ and $B \neq A, B \neq \varnothing$, which means that not every element of A belongs to B. For example above B is a proper subset of A.

There are several very important sets that have universal notation: $\mathbf{N}, \mathbf{Z}, \mathbf{Q}, \mathbf{R}, \mathbf{C}$ (another universal notation for the same sets is $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$). They are, respectively,

- the set of natural numbers $\mathbf{N} = \{0, 1, 2, \ldots\}$ (some people prefer to start with 1);
- the set of integers $\mathbf{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$;
- the set of rational numbers $\mathbf{Q} = \left\{\frac{p}{q} \colon p, q \in \mathbf{Z},\ q \neq 0\right\}$;
- the set of real numbers $\mathbf{R}$;
- the set of complex numbers $\mathbf{C} = \{x + \mathrm{i}y \colon x, y \in \mathbf{R},\ \mathrm{i}^2 = -1\}$.

You can note that I use a different notation to define sets. In particular, the expression $\{p/q \colon p, q \in \mathbf{Z},\ q \neq 0\}$ means a set of expressions of the form p/q, where (this conditional part is separated by colon, or, sometimes, by a vertical bar) the expressions p and q are taken from the set of integers, and I also require that $q \neq 0$. Note that I do not give you a definition of the set of real numbers, we will use them by relying on our intuitive understanding of elements of $\mathbf{R}$. If you never dealt with complex numbers before — do not worry, I will talk about them separately.

In mathematics in order to introduce a more complicated structure than merely a set it is usual to take a set and require that something else must be true for the elements of this set. For example,

Math 329: *Intermediate Linear Algebra* by Artem Novozhilov©
e-mail: artem.novozhilov@ndsu.edu. Spring 2017

[1]Well, to confuse you: Actually we can define what a set is, but this will require a lot of time and not that useful in mathematics as a whole and not useful at all for our course.

very loosely, if I am allowed to add and multiply any two elements of a given set, and I assume that additive and multiplicative inverses (except for 0 element) exist for any elements then I call such set with two additional operations *a field*. I understand that this is a very vague description, but do not worry about it at this point. The key fact that you must remember is that the set of real numbers $\mathbf{R}$ is a field. Therefore at least for the next month if I say "field $\mathbf{F}$" you can replace it with "the set of real numbers $\mathbf{R}$."

Why $\mathbf{R}$ is a field? Take any two real numbers $x, y \in \mathbf{R}$. Then we know that $x + y \in \mathbf{R}$, $xy \in \mathbf{R}$, there is, e.g., $-x$ such that $x + (-x) = 0$ and y^{-1} such that $yy^{-1} = 1$, there is numbers 0 and 1 which have properties that $x + 0 = x$, $1y = y$, moreover, both the addition and multiplication are associative and commutative, moreover, one is distributive with respect to the other.

Exercise 1. Convince yourself that $\mathbf{Q}$ is a field, but neither $\mathbf{Z}$ nor $\mathbf{N}$ is a field.

Now I am prepared to give a definition of the most important in this course mathematical structure — vector (or linear) space.

7.2 Definition of a vector (or linear) space

Definition 7.1. *A vector space over a field* $\mathbf{F}$ *is a nonempty set* V *with two operations: addition and multiplication by scalars from* $\mathbf{F}$*. That is, for any* $u, v \in V$ *and* $\alpha \in \mathbf{F}$

$$u + v \in V,$$
$$\alpha u \in V.$$

The addition and multiplication by scalars satisfy the following axioms of a vector space:

1. *Addition is commutative:* $u + v = v + u$ *for all* $u, v \in V$*;*
2. *Addition is associative:* $u + (v + w) = (u + v) + w$ *for all* $u, v, w \in V$*;*
3. *There is a special vector* $0 \in V$ *such that* $0 + v = v$ *for all* $v \in V$*;*
4. *There is a special vector* $-v \in V$ *such that* $v + (-v) = 0$ *for all* $v \in V$*;*
5. *The multiplicative identity* $1 \in \mathbf{F}$ *satisfies* $1v = v$ *for all* $v \in V$*;*
6. *For all* $\alpha, \beta \in \mathbf{F}$ $(\alpha\beta)v = \alpha(\beta v)$ *for all* $v \in V$*;*
7. *Multiplication by scalars is distributive with respect to vector addition:* $\alpha(u + v) = \alpha u + \alpha v$ *for all* $\alpha \in \mathbf{F}$ *and* $u, v \in V$*;*
8. *Addition of scalars is distributive with respect to scalar multiplication:* $(\alpha_1 + \alpha_2)v = \alpha_1 v + \alpha_2 v$ *for all* $\alpha_1, \alpha_2 \in \mathbf{F}$ *and* $v \in V$*.*

The elements of a vector space are called vectors.

Remark 7.2. Again, as I said above, replace the words "field $\mathbf{F}$" with "set of real numbers $\mathbf{R}$" if you do not feel comfortable talking about abstract fields.

Note that vector space is a set V plus a field. The same set with different fields can give rise to very different vector spaces, and therefore the correct notation for a vector space is $(V, \mathbf{F}, +, \cdot)$, emphasizing that one needs a clear understanding of the nature of set V, field $\mathbf{F}$ and specifics of operations $+$ and $\cdot$. In most books, however, a slight abuse of notations is used, people just write "a vector field V" if everything else is clear from the context. Words "real vector space V" mean a vector space over the field of real numbers. I will use the same conventions in these lectures.

Now when we have a definition of a vector space let us look at some examples.

Example 7.3. The set of real numbers is a real vector space $(\mathbf{R}, \mathbf{R}, +, \cdot)$, where the operations of addition and multiplication are the usual operations.

(To give you a little headache: vector space $(\mathbf{R}, \mathbf{Q}, +, \cdot)$ over the field of rational numbers satisfies all the axioms above, but is fundamentally different from real vector space $\mathbf{R}$. Check the axioms.)

Example 7.4. Recall that Descartes's product $A \times B$ of two sets $A = \{a_1, a_2, \ldots\}$ and $B = \{b_1, b_2, \ldots\}$ is the set $A \times B$ of all possible ordered pairs, i.e.,

$$A \times B = \{(a, b) \colon a \in A,\, b \in B\}.$$

This immediately can be generalized to $A \times B \times \ldots \times Z$. There is a standard notation $A \times A \times \ldots \times A = A^n$ for n times product.

Hence for any field $\mathbf{F}$ I can form the set $\mathbf{F}^n$. In particular $\mathbf{R}^n$ is the set of all ordered n-tuples $(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$, which we so got used to call *vectors*. But please note that this is just one (although arguably most important) example of a vector space. Summarizing, $\mathbf{R}^n$ is a real vector space with the operations of addition and multiplication defined componentwise: For all $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$ $\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$, and for $\alpha \in \mathbf{R}$ $\alpha \boldsymbol{x} = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$. Note that for the vectors in $\mathbf{R}^n$ I will always use the bold font notation $\boldsymbol{x}, \boldsymbol{y}$, consistent with the notation from the first part of the course. Checking the rest of the axioms is left as an exercise.

Example 7.5. As a particular case of the previous example consider $\mathbf{R}^3$, which is our usual Euclidian vector space in which we live. Each vector $(x_1, x_2, x_2) \in \mathbf{R}^3$ can be readily visualized as an arrow from the origin to the point with coordinates (x_1, x_2, x_3) and addition of two vectors and multiplication by a scalar have simple geometric meaning. While it is very useful to keep this geometric picture in mind even if we are talking about other vector spaces, remember that, given the definition, I consider all the vector spaces as algebraic objects, and so almost all of the proofs will be of an algebraic nature.

Example 7.6. The set of matrices $\mathbf{M}_{m \times n}(\mathbf{R})$ with real entries over the reals is a real vector space with the addition and multiplication by scalars defined in the standard way. Please check the axioms.

Example 7.7. Recall that the function f defined as $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$, where I assume that all $a_i \in \mathbf{R}$, is called a *polynomial* of degree n. To be totally precise I need also specify the domain of f, let me take it also $\mathbf{R}$, hence my function becomes a function from reals to reals. I claim that the set of all polynomials over the reals is a vector space, which I denote $\mathbf{P}$. I define the addition of two polynomials $f, g \in \mathbf{P}$ as a function $h = f + g$, which is defined as

$$h(x) = (f + g)(x) = f(x) + g(x).$$

Similarly, for any $\alpha \in \mathbf{R}$, $p = \alpha f$ if

$$p(x) = \alpha f(x).$$

With these definitions one can check that all the axioms hold for my set and hence it is a real vector space.

Example 7.8. Let $\boldsymbol{A}$ be an $m \times n$ matrix with real entries and $\boldsymbol{x} \in \mathbf{R}^n$. Then the set of solutions to the homogeneous system of linear algebraic equations

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

is a vector space. Indeed, let $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ be 2 solutions, then, by the distributive property of matrix multiplication

$$\boldsymbol{A}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \boldsymbol{A}\boldsymbol{x}_1 + \boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0},$$

and hence their sum is also a solution. Similarly, for any $\alpha \in \mathbf{R}$

$$\boldsymbol{A}(\alpha\boldsymbol{x}_1) = \alpha(\boldsymbol{A}\boldsymbol{x}_1) = \alpha\boldsymbol{0} = \boldsymbol{0},$$

and therefore $\alpha\boldsymbol{x}_1$ is also a solution. All axioms hold since each solution is in $\mathbf{R}^n$ (think this out!), but also note that generally not every $\boldsymbol{y} \in \mathbf{R}^n$ is a solution, and hence in general the set of solutions is a *proper* subset of $\mathbf{R}^n$. Such subsets, which also satisfy the definition of a vector space are, quite naturally, called *subspaces*.

In general, I can introduce an important notion of a subspace as follows.

Definition 7.9. *A subspace W of a vector space V over $\mathbf{F}$ is a nonempty subset of V that is closed under the operations of addition and multiplication by scalars.*

Clearly a subspace is a vector space itself (why?). I will denote that W is a subspace of V as $W \subseteq V$.

Example 7.10. Convince yourself that the set of all symmetric matrices $n \times n$ is a subspace of $\mathbf{M}_{n\times n}$.

A subspace that is different from V itself and $\{0\}$ is called *proper*.

Exercise 2. What are the proper subspaces of $\mathbf{R}^2$?

The axioms allow a number of immediate consequences, which are so intuitively true for the familiar real numbers that sometimes it is not clear why we should worry to prove them at all. However, keep in mind that now we are talking about abstract vector spaces, of which $\mathbf{R}$ is just one example, and do not confuse the introduces objects $0, -v$ with the usual number 0 and additive inverse. In proving the following facts one can use only the axioms of a vector space, so it is quite a good exercise in logic and practicing proofs.

Proposition 7.11.

1. *The zero vector $0 \in V$ is unique.*
2. *The additive inverse $-v \in V$ is unique.*
3. *For $0 \in V$ and $\alpha \in \mathbf{F}$ I have $\alpha 0 = 0$. Note that on the right hand side $0 \in V$.*
4. *For $0 \in \mathbf{F}$ and $v \in V$ I have $0v = 0$. Note that on the right hand side $0 \in V$.*
5. *If $\alpha \neq 0$ and $\alpha v = 0$ then $v = 0$.*
6. *For all $v \in V$ I have $-v = (-1)v$.*

Proof. I will prove only the first one, leaving the rest as an exercise.

Assume that there are two such vectors $0, 0'$ that satisfy axiom 3. I have, by this axiom that

$$0 + 0' = 0' \quad 0' + 0 = 0.$$

By axiom 1 $0 + 0' = 0' + 0$ and hence $0 = 0'$.

∎

The associativity of vector addition allows not to use the parenthesis in the expressions like

$$v_1 + (v_2 + v_3) + v_4 = ((v_1 + v_2) + v_3) + v_4 = v_1 + v_2 + v_3 + v_4.$$

I can use this property to define a central notion of a *linear combination*:

Definition 7.12. *Let $v_1, \ldots, v_k \in V$ be a collection of vectors from vector space V (note the abuse of notations here!). A linear combination of $v_1, \ldots, v_k$ is an expression of the form*

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k = \sum_{i=1}^{k} \alpha_i v_i \in V, \quad \alpha_i \in \mathbf{F}.$$

It is very important that the number of vectors in the linear combination is finite.

Note that I gave the definition of a linear combination as a *collection* of vectors, not as a set (I will use also the word *list* as a synonym of *collection*). This allows me to have several identical copies of the same vector in this collection (in a set all the elements are distinct). At some point I will have to also introduce an *ordered* collection (or ordered list) of vectors, where I will have to take into account also the order of my elements. I do not need to do this now.

Definition 7.13. *The set of all vectors that are linear combinations of a collection $S = (v_1, \ldots, v_n)$ is called the span of S and denoted* $\operatorname{span} S$.

Proposition 7.14. *Let $v_1, \ldots, v_k \in V$. Then $\operatorname{span} S = \operatorname{span}(v_1, \ldots, v_k)$ is a subspace of V. Moreover, if $S \subseteq W$, where $W \subseteq V$, then $\operatorname{span} S \subseteq W$.*

Exercise 3. Prove the proposition.

Definition 7.15. *A collection S of vectors $v_1, \ldots, v_k$ is called linearly independent if the linear combination*

$$\alpha_1 v_1 + \ldots \alpha_k v_k, \quad \alpha_i \in \mathbf{F}$$

is equal to zero if and only if $\alpha_1 = \ldots = \alpha_k = 0$.

A collection of vectors is called linearly dependent if there are scalars α_i, $i = 1, \ldots, k$, not equal zero simultaneously, such that

$$\alpha_1 v_1 + \ldots \alpha_k v_k = 0, \quad \alpha_i \in \mathbf{F}.$$

Note that any list that contains zero vector or two identical vectors is linearly dependent (why?). A list that contains only one vector is linearly independent only if this vector is different from zero. Similarly, (v_1, v_2) is linearly independent if none of the vectors is a multiple of another.

Here is another restatement of the property to be linearly dependent.

Proposition 7.16. *A collection S is linearly dependent if and only if one of the vectors in S can be represented as a linear combination of the rest.*

Exercise 4. Prove the proposition.

How to determine whether a given list is linearly dependent or independent? It is especially simple in the case of $\mathbf{R}^n$. So consider $S = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ is given. Here I treat each vector as a column vector. I form a linear combination and equal it to zero:

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

In more details, denoting v_{ij} the i-th coordinate of the j-th vector $\boldsymbol{v}_j$:

$$\alpha_1 v_{i1} + \ldots + \alpha_k v_{ik} = 0, \quad i = 1, \ldots, n.$$

In other words, this is a system of linear algebraic equations with the matrix $\boldsymbol{A} = [\boldsymbol{v}_1 \mid \ldots \mid \boldsymbol{v}_k]$ and the vector of unknowns $\boldsymbol{\alpha} = [\alpha_1 \ \ \ldots \ \ \alpha_k]^\top$, which I can write as

$$\boldsymbol{A}\boldsymbol{\alpha} = \mathbf{0}.$$

This system always has a trivial solution. If there is a nontrivial solution this would imply that my list is linearly dependent. Otherwise, it is linearly independent. In particular, using obtained earlier results, I conclude that if $k > n$, that is, if the number of variables in my system is bigger than the number of equations, my list is *always* linearly dependent because I always have a nontrivial solution.

I can summarize the discussion above as follows.

Proposition 7.17. *Let $S = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ be a collection of vectors from $\mathbf{R}^n$. Then this collection is linearly independent if the row reduced echelon form of matrix $\boldsymbol{A} = [\boldsymbol{v}_1 \mid \ldots \mid \boldsymbol{v}_k]$ (i.e., the matrix composed of vectors from S as columns) has no free variables.*

Example 7.18. Consider $\mathbf{R}^3$ and the list of four vectors

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad \boldsymbol{v}_4 = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}.$$

From the previous I immediately conclude that $S = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4)$ is linearly dependent. What about $S' = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$? I have now the system

$$\boldsymbol{A}\boldsymbol{\alpha} = \mathbf{0}, \quad \boldsymbol{A} = [\boldsymbol{v}_1 \mid \boldsymbol{v}_2 \mid \boldsymbol{v}_3].$$

The matrix is invertible (its determinant is not zero), and hence this list is linearly independent.

Now, having at my disposal the notions of a span and linearly independent set I can define a basis of a vector space.

Definition 7.19. *A basis of a vector space V is a collection $\mathcal{B}$ of vectors that is both linearly independent and spans V.*

Example 7.20. Consider the standard unit vectors $\boldsymbol{e}_i \in \mathbf{R}^n$, i.e., vectors with all entries equal 0 except for the i-th coordinate, which is equal to 1. I claim that $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ is a basis of $\mathbf{R}^n$. Indeed, from the previous, my matrix $\boldsymbol{A} = [\boldsymbol{e}_1 \mid \ldots \mid \boldsymbol{e}_n] = \boldsymbol{I}$ and hence the system $\boldsymbol{A\alpha} = \boldsymbol{0}$ has only the trivial solution and hence this list is linearly independent. For any other vector $\boldsymbol{x} = [x_1 \ \ldots \ x_n]^\top \in \mathbf{R}^n$ I have

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n,$$

and therefore my vectors span $\mathbf{R}^n$.

Generalizing, any list of vectors $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is a basis of $\mathbf{R}^n$ if the matrix $\boldsymbol{A} = [\boldsymbol{v}_1 \mid \ldots \mid \boldsymbol{v}_n]$ is invertible, which is probably easiest to check by calculating the determinant.

Proposition 7.21. *Collection $\mathcal{B} = (v_1, \ldots, v_n)$ is a basis of V if and only if any vector $w \in V$ can be written in a unique way as a linear combination*

$$w = \alpha_1 v_1 + \ldots + \alpha_n v_n, \quad \alpha_i \in \mathbf{F}.$$

Proof. ($\Longrightarrow$). Assume that $\mathcal{B}$ is a basis. I need to show that any vector can be represented in a unique way as a linear combination of vectors from $\mathcal{B}$. Assume, looking for a contradiction, that vector $w \in V$ can be represented in two different ways:

$$w = \alpha_1 v_1 + \ldots + \alpha_n v_n,$$
$$w = \alpha_1' v_1 + \ldots + \alpha_n' v_n.$$

By subtracting these two equalities I get

$$0 = (\alpha_1 - \alpha_1')v_1 + \ldots + (\alpha_n - \alpha_n')v_n.$$

Since $\mathcal{B}$ is linearly independent hence the expression above is equal to zero only if $\alpha_i - \alpha_i' = 0$ for all i, which implies $\alpha_i = \alpha_i'$.

($\Longleftarrow$) Assume that any vector $w \in V$ can be represented as a linear combination of vectors from $\mathcal{B}$ in a unique way. This also must be true, hence, for $w = 0$:

$$0 = \alpha_1 v_1 + \ldots + \alpha_n v_n.$$

Since this representation is unique, all α_i must be zero, otherwise I could have multiplied by $\beta \neq 1$ and got another linear combination. Hence $\mathcal{B}$ is linearly independent. The assumption that $\mathcal{B}$ spans V is hidden in the initial assumption, and hence $\mathcal{B}$ is a basis. ∎

Since the scalars in the representation by a linear combination of basis vectors are unique, I can can call them the *coordinates* of vector $w \in V$ with respect to the basis $\mathcal{B}$. Note that for different bases I will have different coordinates. In particular, basis $\mathcal{B} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ is called the standard basis of $\mathbf{R}^n$ for the reason when we say that vector $\boldsymbol{x}$ is $[1 \ 2 \ 3]^\top$, we actually mean that these are the coordinates of $\boldsymbol{x}$ with respect to the standard basis.

Here is another basic fact, which I will leave as one more proof exercise.

Proposition 7.22. *Let S be a list of vectors from V, and $w \in V$, and let $S' = (S, w)$, i.e., the list obtained by adding w to S. Then $\operatorname{span} S = \operatorname{span} S'$ if and only if $w \in \operatorname{span} S$. Assume that S is linearly independent, then S' is independent if and only if $w \notin \operatorname{span} S$.*

Now let me significantly reduce the number of vector spaces, which we will be studying. In particular, I will consider only *finite-dimensional* vector spaces.

Definition 7.23. *If there is a finite list S that spans V then V is called finite dimensional. In particular, any vector space with a basis*[2] *is finite dimensional. Vector spaces that are not finite dimensional are called infinite dimensional.*

We will not study infinite dimensional vector spaces, this mostly belongs to the mathematical subfield called *functional analysis*.

Exercise 5. Show that $(\mathbf{R}, \mathbf{Q}, +, \cdot)$ is an infinite dimensional vector space.

Exercise 6. Show that the vector space of all polynomials is infinite dimensional.

Proposition 7.24. *Let V be a finite dimensional vector space. Let S be a finite subset that spans V and let R be a linearly independent subset of V. Then one can obtain a basis by adding elements from S to R.*

Proof. If $S \subseteq \operatorname{span} R$ then R is a basis by definition. Assume that S is not a subset of $\operatorname{span} R$. Then there is an element $s \in S$ that is not in $\operatorname{span} R$, form $R' = (R, s)$. By proposition 7.22 R' is independent. Since S is finite we can only add a finitely many new vectors. Since S spans V we eventually will have a basis. ∎

Proposition 7.25. *Let V be a finite dimensional vector space. Let S be a finite subset that spans V. One can obtain a basis of V by deleting elements from S.*

Proof. If S is independent we are done. If S is not independent, there is a vector $s \in S$ which is a linear combination of the rest of the vectors in S. Form S' by dropping s. By Proposition 7.22, $\operatorname{span} S' = \operatorname{span} S$ and hence S' spans V. Since S is finite, this process will end at some point and we will have a basis. ∎

Remark 7.26. For those who likes mathematical nitpicking: What is one problem with the proof above?

Now we are ready to prove the main result of this introductory lecture on vector spaces.

Theorem 7.27. *Let S and R be finite subsets of vector space V. Assume that $V = \operatorname{span} S$ and R is linearly independent. Then then the number of elements in S, which is denoted $|S|$, is not less than the number of elements in R, which is denoted $|R|$:*

$$|S| \geq |R|.$$

Proof. I will prove this theorem by contradiction. In particular, I will assume that $|S| < |R|$ and deduce that in this case R must be dependent.

Say, $S = (s_1, \ldots, s_m)$ and $R = (r_1, \ldots, r_n)$, and $m < n$ by assumption. Since S spans V, each element of V can be represented as a linear combination of vectors in S, in particular,

$$r_1 = \alpha_1 s_1 + \ldots + \alpha_m s_m.$$

[2] "...with a basis" how I defined it! There are different, more general definitions of a basis.

Assume without loss of generality that $\alpha_1 \neq 0$ (we must have at least one non-zero α_i and always can call it α_1 after reindexing the elements). Hence I can write that

$$s_1 = \alpha_1^{-1} r_1 - \alpha^{-1}\alpha_2 s_2 - \ldots - \alpha^{-1}\alpha_m s_n.$$

Since, by the above, $s_1 \in \text{span}(r_1, s_2, \ldots, s_m)$ then $V = \text{span}(r_1, s_2, \ldots, s_m)$. The idea is keep replacing $s_2, s_3, \ldots$ one by one with $r_2, r_3, \ldots$. For example, since $r_2 \in V$, I have

$$r_2 = \beta_1 r_1 + \beta_2 s_2 + \ldots + \beta_m s_m.$$

Moreover, I know that at least one of the constants $\beta_2, \ldots, b_m$ is non-zero, since, if I assume that they all zero, then I would have a linear relation between the elements of a linearly independent list, a contradiction. I can reorder my elements in such a way so that non-zero β_2 corresponds to s_2, and hence I have

$$s_2 = \beta_2^{-1} r_2 - \beta_2^{-1}\beta_1 r_1 - \ldots - \beta_2^{-1}\beta_m s_m.$$

This means that now I have the list $(r_1, r_2, s_3, \ldots, s_m)$, which spans V. I can continue this process (an accurate argument requires induction) and get at the last step that $(r_1, \ldots, r_m)$ spans V, but since $m < n$, hence there should be some vectors left in the list R, which can be represented as linear combinations of $(r_1, \ldots, r_m)$, therefore R is not linearly independent, I get a contradiction and conclude that $m \geq n$. ■

Theorem 7.28. *Let V be a finite dimensional vector space. Then any two bases of V have the same finite number of elements.*

Proof. Since by definition there is a finite spanning list, and by the previous theorem the number of elements in this list must not be less than the number of elements in a basis, therefore any basis is finite. Now, let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two bases. By the previous theorem $|\mathcal{B}_1| \leq |\mathcal{B}_2|$ and $|\mathcal{B}_2| \leq |\mathcal{B}_1|$. Therefore,

$$|\mathcal{B}_1| = |\mathcal{B}_2|.$$

■

So, concluding and summarizing several results: Any basis of a finite dimensional vector space has the same number of vectors; any spanning list has not fewer elements than a basis; any linearly independent list has no more elements than a basis. A spanning list is a basis if and only if it has the same number of elements as a basis; linearly independent list is a basis if and only if it has the same number of elements as a basis.

Exercise 7. Prove what has not been proved in the last paragraph.

Finally,

Definition 7.29. *The dimension,* $\dim V$*, of a finite dimensional vector space V is the number of vectors in a basis.*

Proposition 7.30. *If W is a subspace of a finite dimensional vector space V, then W is finite dimensional, and* $\dim W \leq \dim V$*. Moreover,* $\dim W = \dim V$ *if and only if $W = V$.*

Exercise 8. Prove the last proposition.

Exercise 9. Let $W_1 \subseteq V$ and $W_2 \subseteq V$ be two subspaces of a vector space V. Show that their intersection $W_1 \bigcap W_2$, that is, the set of all the vectors that belong both to W_1 and W_2, is also a subspace. Show that their union is not necessarily a subspace.

Exercise 10. Let $W_1 \subseteq V$ and $W_2 \subseteq V$ be two subspaces of a vector space V. I define $W_1 + W_2$ as the set of all possible vectors of the form

$$w_1 + w_2,$$

where $w_1 \in W_1$ and $w_2 \in W_2$. Show that $W_1 + W_2$ is a subspace of V.

Exercise 11. Let $W_1 + W_2 = V$ and $W_1 \bigcap W_2 = \{0\}$. Prove that any vector $v \in V$ can be *uniquely* represented as

$$v = w_1 + w_2,$$

where $w_1 \in W_1$, $w_2 \in W_2$.

Exercise 12. Let $W_1 \subseteq V$ and $W_2 \subseteq V$ be two subspaces of a finite dimensional vector space V. Prove that

$$\dim W_1 + \dim W_2 = \dim(W_1 \bigcap W_2) + \dim(W_1 + W_2).$$