ETMAG

LECTURE 12

- Vector spaces, cont'd
- Subspaces
- Linear independence

Example. (A REALLY outlandish one)

Let X be any set. We will use $V = (2^X, \div)$ as the Abelian group of vectors, where $\div$ denotes the operation of symmetric difference of sets, $A \div B = (A \cup B) \setminus (A \cap B)$. We will also use $(\mathbb{Z}_2, \oplus, \otimes)$ as the field of scalars. Scaling is defined as follows:

for every set A, $0 \cdot A = \varnothing$ and $1 \cdot A = A$.

Comprehension.
Check that $(V, \mathbb{Z}_2, \cdot)$ is a vector space.

FAQ. 1

What the hell is a vector?
The only proper answer to this question, even though a little tautological, is *"A vector is an element of a vector space"*. The previous example teaches us that sets can be vectors. In other examples we have seen numbers, complex numbers, n-tuples of numbers, functions, polynomials etc. playing the role of vectors.

FAQ. 2
What the hell is a scalar then?

Well, you probably realize that the answer will be equally trivial (or disturbing). An element of a field $\mathbb{K}$ may be called a scalar if somebody decides to construct a vector space using $\mathbb{K}$ as the second element of the ordered triple constituting a vector space. In particular, if we consider $\mathbb{K}$ a vector space over itself then all elements of $\mathbb{K}$ are at the same time scalars and vectors.

Example.
In the vector space $\mathbb{R}$ over the field $\mathbb{R}$, real numbers are both vectors and scalars.

In $\mathbb{C}$ over $\mathbb{R}$ complex numbers are vectors, real numbers are scalars.

In 2^X over $\mathbb{Z}_2$vectors are subsets of X and there are but two scalars, 0 and 1.

What makes study of general vector spaces useful is that whatever facts we discover about vector spaces in general they are true in each of these spaces.

For example, we gave the following theorem:

Theorem. (Arithmetic properties of vector spaces)
In every vector space V over a field $\mathbb{K}$

1. for every vector v, $0 \cdot v = \Theta$, (Θ is the zero vector, 0 is the zero scalar).
2. for every scalar p, $p \cdot \Theta = \Theta$.
3. for every scalar p and for every vector v, $(-p) \cdot v = p \cdot (-v) = -(p \cdot v)$.
4. for every scalar p and for every vector v, $p \cdot v = \Theta$ implies $p = 0$ or $v = \Theta$.

Comprehension. (Prove the theorem).

Hint. $0 \cdot v = (0 + 0) \cdot v$

Definition.

Let V be a vector space over $\mathbb{K}$. A subset $W \subseteq V$ is called a *subspace* of V if W is a vector space over $\mathbb{K}$ under "the same" operations of vector addition and scalar multiplication.

Theorem.

W is a subspace of V if and only if

1. $(\forall \alpha \in \mathbb{K})(\forall w \in W)\ \alpha w \in W$ (W is "*closed under scaling*")
2. $(\forall w_1, w_2 \in W)\ w_1 + w_2 \in W$ (W is "*closed under vector addition*")
3. $W \neq \emptyset$ (or, equivalently $\Theta_V \in W$) (W is *nonempty, contains* Θ).

Proof. (Outline)

The "only if" part is trivial.

The "if" part: 1.,2. and 3., together with last lecture theorem imply that $(W, +)$ is a subgroup of $(V, +)$ and that it is closed under scaling. The remaining axioms of a vector space follow from the simple fact that all vectors of W belong to V, which means they satisfy all required identities. QED

Comprehension. (subspaces)
Decide which of the following subsets are subspaces:

1. $\{(x, y) \in \mathbb{R}^2 : xy \geq 0\}$ in $\mathbb{R}^2$ over $\mathbb{R}$
2. $\{(x, y) \in \mathbb{R}^2 : x + y \geq 0\}$ in $\mathbb{R}^2$ over $\mathbb{R}$
3. $\{(x, y) \in \mathbb{R}^2 : x = 5y\}$ in $\mathbb{R}^2$ over $\mathbb{R}$
4. $\{(x, y) \in \mathbb{R}^2 : x^2 = y\}$ in $\mathbb{R}^2$ over $\mathbb{R}$
5. $\{(x, y, z) \in \mathbb{R}^3 : x + y - 3z = 1\}$ in $\mathbb{R}^3$ over $\mathbb{R}$
6. $\left\{\{a, b\}, \{a\}, \emptyset\right\}$ in $2^{\{a,b,c\}}$ over $\mathbb{Z}_2$
7. The set of all finite sets from $2^{\mathbb{N}}$ over $\mathbb{Z}_2$

Comprehension self-test.

Find all subspaces of $\mathbb{R}^2$ over $\mathbb{R}$ (with usual operations).

Definition.

Let V be a vector space over a field $\mathbb{K}$, let and let $v_1, \dots, v_n \in V$. A vector v is called a *linear combination* of vectors $v_1, \dots, v_n$ iff there exist $a_1, \dots, a_n \in \mathbb{K}$ such that $v = a_1 v_1 + \cdots + a_n v_n$.

We also use the sigma notation, $v = \sum_{s=1}^{n} a_s v_s$.

We say that v is **the** *linear combination* of vectors $v_1, \dots, v_n$ with coefficients $a_1, \dots, a_n$.

A common problem in linear algebra is to decide whether a given vector is or is not a linear combination of other given vectors.

Example.

If you fail ETMAG you might decide to blow-up Polytechnica in revenge. A recipe found on the darknet says that mixing 30% of ingredient A, 50% of B and 20% of C will provide an explosive. A leading branch of toothpaste T consists of 10, 60 and 30 percent of those, a scouring powder S has 5, 80 and 15 and a washing machine powder P has 25, 50 and 25. Can you get your explosive mixing those in some proportion?

In other words, do there exist coefficients t, s, p such that

$$(30,50,20) = t(10,60,30) + s(5,80,15) + p(25,50,25)$$

i.e., is (30,50,20) is a linear combination of (10,60,30), (5,80,15) and (25,50,25).

Note. *In this example we must also require that all coefficients are* ≥ 0.

Comparing component-by-component the left-hand side of

(30,50,20) = *t*(10,60,30) + *s*(5,80,15) + *p*(25,50,25)

to the right-hand side we get a system of equations

$$(*)\begin{cases} 30 = 10t + 5s + 25p \\ 50 = 60t + 80s + 50p \\ 20 = 30t + 15s + 25p \end{cases}$$

We can phrase our problem as:

Does the vector (30,50,20) belong to the set of all possible linear combinations of (10,60,30), (5,80,15) and (25,50,25)?

Or:

Is the system of equations $(*)$ solvable?

Definition.

Let V be a vector space over a field $\mathbb{K}$ and let $S \in V$ be a set of vectors. The *span of the set S* is the set $span(S) \subseteq V$ defined by

$span(S) = \{a_1 v_1 \ + \ ... + \ a_n v_n | n \in \mathbb{N} \wedge (\forall i)(a_i \in \mathbb{K} \wedge \ v_i \in S)\}$ if $S \neq \emptyset$ and

$span(S) = \{\Theta\}$ if $S = \emptyset$.

In plain language, $span(S)$ is the set of all possible linear combinations of vectors from S.

Example.

Consider $\mathbb{R}^3$ and two vectors $(1, 2, 0), (2, 4, 0) \in \mathbb{R}^3$.

$span(\{(1, 2, 0), (0, 0, 3)\}) = \{a(1, 2, 0) + b(2, 4, 0) | a, b \in \mathbb{R}\} =$

$\{(a + 2b, 2a + 4b, 0) | \ a, b \in \mathbb{R}\} = \{(a + 2b, 2(a + 2b), 0) | \ a, b \in \mathbb{R}\} = \{c(1,2,0) | c \in \mathbb{R}\} = \{(x, y, z) | y = 2x \ \& \ z = 0\}$. This is the line passing thorough the origin and $A(1,2,0)$. It is a subspace in $\mathbb{R}^3$.

Theorem.

Let V be a vector space over a field $\mathbb{K}$ and let S be a subset of V. Then $\mathrm{W} = span(S)$ is a subspace of V.

Proof.

If S is empty, $span(S) = \{\Theta\}$ – a subspace of V.

If $S \neq \emptyset$ and $u, v \in span(S)$, we can choose vectors $v_1, \dots, v_n \in S$ so that $u = a_1 v_1 + \cdots + a_n v_n$ and $v = b_1 v_1 + \dots + b_n v_n$ for some $a_i, b_i \in \mathbb{K}$. Clearly, $u + v = a_1 v_1 + \dots + a_n v_n + b_1 v_1 + \dots + b_n v_n = (a_1 + b_1) v_1 + \dots + (a_n + b_n) v_n \in W$. For every scalar p, $pu = p(a_1 v_1 + \dots + a_n v_n) = (pa_1) v_1 + \dots + (pa_n) v_n \in W$. QED

Theorem. (Alternate definition of *span*)
Let $S \subseteq V$. Then $span(S)$ is the smallest subspace of V containing S.
Proof.
Every subspace of V containing S must contain all linear combinations of vectors from S. QED

We call $span(S)$ the *subspace* (*of* V) *spanned by S.*

One advantage of the alternate definition over the other one is that it covers the case $S = \emptyset$ without branching.

Fact.

Let $V(S)$ denote the set of all subspaces of V containing S. Then

$$span(S) = \bigcap_{T \in V(S)} T$$

Proof. It is enough to show that intersection of a collection of subspaces is a subspace of V, which is easy. (Each contains Θ so the intersection does, too, etc.). QED

Examples.

In $\mathbb{C}$ over $\mathbb{R}$:

$span(\{1, i\}) = \{a \cdot 1 + b \cdot i | a, b \in \mathbb{R}\} = \mathbb{C}$

$span(\{1 + i\ , 2\ +\ i\})\ = \mathbb{C}$

$span(\{i\ + 2, 2i + 4\}) = \{2a\ +\ ai \colon\ a \in \mathbb{R}\} \neq \mathbb{C}$

In $\mathbb{R}[x]$ over $\mathbb{R}$:

$span(\{x^2, x, 1\}) = \{ax^2 + bx + c | a, b, c \in \mathbb{R}\}$ – the set of all polynomials of degree at most 2.

In 2^X over $\mathbb{Z}_2$:

$span(\{A, B\}) = \{\emptyset, A, B, A \div B\}$ because $A \div (A \div B) = (A \div A) \div B = \emptyset \div \mathrm{B} = B$.

Theorem (Properties of *span*)

Let V be a vector space over a field $\mathbb{K}$ and let $S, T \subseteq \text{V}$. Then

1. $S \subseteq span(S)$
2. $span\big(span(S)\big) = span(S)$
3. $S \subseteq T \Rightarrow span(S) \subseteq span(T)$
4. $(\forall v \in V)(v \in span(S) \Leftrightarrow span(S) = span(S \cup \{v\})$

Proof.

1. 2. and 3. are rather obvious.

Proof of 4.

$(\forall v \in V)(v \in span(S) \Leftrightarrow span(S) = span(S \cup \{v\})$

($\Rightarrow$) From 3., $span(S) \subseteq span(S \cup \{v\})$. Suppose $w \in span(S \cup \{v\})$ i.e., $w = av + b_1u_1 + \cdots + b_ku_k$ for some $u_1, \dots, u_k \in S$ and some scalars $b_1, \dots, b_k$. $v \in span(S)$ means $v = c_1v_1 + \cdots + c_nv_n$ for some $v_1, \dots, v_n \in S$ and $c_1, \dots, c_n \in \mathbb{K}$. Hence, $w = a(c_1v_1 + \cdots + c_nv_n) + b_1u_1 + \cdots + b_ku_k = (ac_1)v_1 + \cdots + (ac_n)v_n + b_1u_1 + \cdots + b_ku_k \in span(S)$.

($\Leftarrow$) $span(S) = span(S \cup \{v\})$ means that every vector from $span(S \cup \{v\})$, including v, belongs to $span(S)$. QED

Definition.

Let V be a vector space over a field $\mathbb{K}$ and let $S = \{v_1, \dots, v_n\}$ be a set of vectors from V. S is said to be *linearly independent* iff

$$(\forall a_1, a_2, \dots, a_n \in \mathbb{K})(a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = \Theta \Rightarrow a_1 = a_2 = \cdots = a_n = 0)$$

If S is not linearly independent, S is said to be *linearly dependent*.

Remark.
This definition may be confusing: it is NOT that the linear combination with all zero coefficients is Θ (which is trivially true for any set of vectors, linearly independent or not) but the other way around, the condition for a set to be linearly independent is that the **only** way to make a linear combination of its vectors Θ is to make all coefficients 0.

Examples.

1. Is $S = \{(1,0,1), (1,1,0), (0,1,1)\}$ linearly independent in $\mathbb{R}^3$?
   Suppose $a(1,0,1) + b(1,1,0) + c(0,1,1) = (0,0,0)$. This means
   $a + b = 0$
   $b + c = 0$
   $a + c = 0.$
   Subtracting the second equation from the first we get $a - c = 0$, i.e., $a = c$. Replacing a with c in the third we get $2c = 0$ hence, $c = 0$. This easily implies that $b = a = 0$. The answer is YES.
2. The empty set $\emptyset$ is linearly independent.

Examples.

3. Is $S = \{(1,0,-1),(1,1,0),(0,1,1)\}$ linearly independent in $\mathbb{R}^3$?
   Suppose $a(1,0,-1) + b(1,1,0) + c(0,1,1) = (0,0,0)$. Then

$$\begin{cases} a + b \quad\quad = 0 \\ \quad\quad b + c = 0 \\ -a + \quad c = 0 \end{cases} (e_1+e_3) \rightarrow \begin{cases} b + c = 0 \\ b + c = 0 \\ -a + c = 0 \end{cases} (e_1-e_2) \rightarrow$$

$$\begin{cases} 0 = 0 \\ b + c = 0 \\ -a + c = 0 \end{cases}$$

   so, $a = c, b = -c$ and no restrictions on c. Putting $c = 1$ we get $a = 1$ and $b = -1$ i.e., we have found non-zero coefficients for our linear combination.
   Conclusion: The set is linearly **dependent**.

Notice: In example 3, $v_1 = v_2 - v_3$. This is no coincidence.

Theorem.

Let V be a vector space over a field $\mathbb{K}$. A set $S = \{v_1, v_2, \dots, v_n\} \subseteq \text{V}$ is linearly independent iff no vector from S is a linear combination of the remaining $n - 1$ vectors.

Proof. (⇒, *by contraposition*)
Suppose a vector from S is a linear combination of the other vectors. *Without loss of generality*, we may assume that v_n is one such, i.e., $v_n = a_1 v_1 + \cdots + a_{n-1} v_{n-1}$. We may write $\Theta = a_1 v_1 + \cdots + a_{n-1} v_{n-1} + (-1) v_n$. In every field $(-1) \neq 0$ hence, the set $\{v_1, v_2, \dots, v_n\}$ is linearly dependent.

Notice:
The phrase "*Without loss of generality*" is used when, instead of considering an arbitrary case (here some v_k), we consider a specific one (here v_n) because, like here,
(a) it makes no difference
(b) it simplifies notation.

Proof.(⇐, also by contraposition)
Suppose $\{v_1, v_2, \dots, v_n\}$ is linearly dependent, i.e., there exist coefficients $a_1, a_2, \dots, a_n$, not all of them zeroes, such that $\Theta = a_1 v_1 + \cdots + a_{n-1} v_{n-1} + a_n v_n$. Without losing generality, we may assume that $a_n \neq 0$ (we can always re-order S so that the vector with the non-zero coefficient is the last one). So

$$a_n v_n = (-a_1) v_1 + \cdots + (-a_{n-1}) v_{n-1}$$

Since a_n, being a nonzero scalar is invertible (w.r.t. multiplication in the field $\mathbb{K}$, we have

$$v_n = (-a_1 a_n^{-1}) v_1 + (-a_2 a_n^{-1}) v_2 + \dots + (-a_{n-1} a_n^{-1}) v_{n-1}.$$ QED