LECTURE 13: DIRECT SUMS AND SPANS OF VECTOR SPACES

MA1111: LINEAR ALGEBRA I, MICHAELMAS 2016

1. Direct sums

Another way to build new vector spaces from old ones is to use **direct sums**. There are two ways to think about this, which are slightly different, but morally the same. First, we define the (external) direct sums of any two vectors spaces V and W over the same field F as the vector space $V \oplus W$ with its set of vectors defined by

$$V \oplus W = V \times W = \{(v, w) : \ v \in V, \ w \in W\}$$

(the $\times$ here is the Cartesian product of sets, if you have seen it, which is defined as a set of ordered pairs as in the second equality on the last line) and with the natural vector operations

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2),$$
$$c(v, w) = (cv, cw).$$

In the last two lines, vector operations take place within V in the first component and in W in the second component.

Example. *The definition we gave for F^2 is just a special case of this definition; that is, $F^2 = F \oplus F$. More generally, F^n is just the n-fold direct product*

$$F \oplus F \oplus \ldots \oplus F.$$

In many very important situations, we start with a vector space V and can identify subspaces "internally" from which the whole space V can be built up using the construction of direct products. While this is slightly untrue, it basically is with respect to any vector space structure we might care about (in more fancy language which we haven't seen yet, the following construction will give an "isomorphic" structure as the above does; for now, think of this as something like if I take all elements of a vector space and paint them green, they may look a little different, but the linear algebra doesn't even notice).

To define this, we first define, for any subsets $X, Y \subseteq V$ their sum

$$X + Y = \{x + y : x \in X, \quad y \in Y\}.$$

Then if W_1, W_2 are subspaces of V with $W_1 \cap W_2 = \{0\}$ (the only vector in both of them is 0) and $W_1 + W_2 = V$, then we say that V is the *(internal) direct sum* of W_1

Date: October 28, 2016.

and W_2, and we write $V = W_1 \oplus W_2$. Again, although this looks slightly different from the definition of direct sums above, we will use the phrase "direct sum" to refer to both; as usual, you must use context to determine which one makes sense! Another way of thinking about this definition, which we may not use much but is nice to think about, is that a vector space V is a direct sum of W_1 and W_2 if and only if every element of V can be written *uniquely* as a sum $w_1 + w_2$ with $w_1 \in W_1$ and $w_2 \in W_2$.

Example. *Consider the plane $P = \{(x, y, z) \in \mathbb{R}^3 : z = 0\}$ in $\mathbb{R}^3$ with $L = \{(x, y, z) \in \mathbb{R}^3 : x = y = 0\}$. Then these are both subspaces of $\mathbb{R}^3$ as we saw above, and their intersection is trivial (as if $z = 0$ and $x = y = 0$, then $x = y = 0$ so $(x, y, z) = 0 \in \mathbb{R}^3$) and their sum is all of $\mathbb{R}^3$ (if $(x, y, z) \in \mathbb{R}^3$, write $(x, y, z) = (x, y, 0) + (0, 0, z)$.), so that*

$$\mathbb{R}^3 = L \oplus P.$$

Example. *Previously, we showed that the set of even polynomials (those with $f(-x) = f(x)$) form a subspace $\mathcal{E}(\mathbb{R})$ of the space of polynomials $\mathcal{P}(\mathbb{R})$. Similarly, the set of odd polynomials $\mathcal{O}(\mathbb{R})$ with $f(-x) = -f(x)$ is a subspace. Moreover, we claim that $\mathcal{P}(\mathbb{R}) = \mathcal{E}(\mathbb{R}) \oplus \mathcal{O}(R)$. To show this, note that if f is both even and odd, then $f(-x) = f(x) = -f(x)$, which means that $f(x) = 0$, and so $\mathcal{E}(\mathbb{R}) \cap \mathcal{O}(\mathbb{R}) = \{0\}$. Moreover, $\mathcal{P}(\mathbb{R}) = \mathcal{E}(\mathbb{R}) + \mathcal{O}(R)$, as for any polynomial f, we can write*

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2},$$

which is a sum of an even polynomial and an odd one.

Example. *Consider the plane $P = \{(x, y, z) \in \mathbb{R}^3 : x - y + 2z = 0\}$ and its normal line $L = \{t(1, -1, 2) : t \in \mathbb{R}\}$. These are both subspaces of $\mathbb{R}^3$, as we have seen. We also claim that $\mathbb{R}^3 = L \oplus P$, as in the geometric example above. To check this, note that a plane and its normal line clearly only intersect in one point, and in this case both pass through the origin, so clearly $L \cap P = \{0\}$. Checking that $\mathbb{R}^3 = L + P$ is the same as checking that for any $(x, y, z) \in \mathbb{R}^3$, there is a point on the line, say $(a, -a, 2a)$, and a a point on the plane, say of the form $(b, b + 2c, c)$, for which*

$$(x, y, z) = (a, -a, 2a) + (b, b + 2c, c) = (a + b, -a + b + 2c, 2a + c),$$

or $x = a + b$, $y = -a + b + 2c$, $z = 2a + c$. This is the same as solving the equation $Ax = d$ with $d = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and $A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix}$. Since we want to show that this is always solvable, for any choice (x, y, z), this is equivalent to A being invertible, or $\det A \neq 0$. Indeed, this example has determinant 6.

N.B. Another way to solve this problem would be to first pick any two non-parallel, non-zero vectors on the plane, and then the second part is really just asking whether the span of those two vectors together with any non-zero vector on the line spans $\mathbb{R}^3$; see the examples in the next section on spans.

2. Spans

Last time, we saw a number of examples of subspaces and a useful theorem to check when an arbitrary subset of a vector space is a subspace. There is one particularly useful way of building examples of subspaces, which we have seen before in the context of systems of linear equations. The idea follows a common theme in mathematics; it is always a good idea to break down a large, complicated set of objects to reduce to the study of a smaller or simpler set of objects. For example, in our computations of determinants, it was extremely useful to reduce the det function to a function on the standard unit basis vectors $e_1, \ldots, e_n$ using multilinearity. The notion we are after is that of linear combination and span. Since we now have a more general notion of vectors, we can recall this setup and generalize it by giving the following definition.

Definition. A vector v in a vector space V over a field F is a **linear combination** of the vectors $v_1, \ldots, v_n \in V$ if there are scalars $c_1, \ldots, c_n \in F$ with

$$c_1v_1 + c_2v_2 + \ldots + c_nv_n = v.$$

Given a set of vectors $v_1, \ldots v_n$, their **span** is the set of linear combinations

$$\mathrm{span}(v_1, \ldots v_n) = \{c_1v_1 + c_2v_2 + \ldots + c_nv_n : c_1, \ldots c_n \in F\}.$$

More generally, given an arbitrary (possibly infinite) set of vectors $S \subset V$, their **span** is the set of all **finite linear combinations** of elements of S:

$$\mathrm{span}(S) = \{c_1v_1 + c_2v_2 + \ldots + c_nv_n : c_1, \ldots c_n \in F, v_1, \ldots, v_n \in S\}.$$

We also say that a set of vectors $S \subseteq$ **spans** or **generates** a vector space V if $V = \mathrm{span}(S)$. In this situation, we also say that S is a **complete** set for V. Finally, we make the convention that if $\emptyset$ is the empty set, then $\mathrm{span}(\emptyset) = \{0\}$.

For us, the reason these are useful is stated in the following result.

Theorem. *For any subset $S \subseteq V$, the set $\mathrm{span}(S)$ is a subspace of V.*

Proof. It is clear that the span of any set contains 0 (note our convention on empty sets). To show that the span is a subspace, we therefore only have to show that it is closed under addition and scalar multiplication. Indeed, if we take two elements of the span to be (note that we assume that the two elements of the span are linear combinations of the exact same set of vectors by possibly adding in some extra factors with 0 coefficients in front)

$$(c_1v_1+c_2v_2+\ldots+c_nv_n)+(c_1'v_1+c_2'v_2+\ldots+c_n'v_n) = (c_1+c_1')v_1+(c_2+c_2')v_2+\ldots+(c_n+c_n')v_n \in \mathrm{span}(S).$$

Similarly,

$$c(c_1v_1 + c_2v_2 + \ldots + c_nv_n) = cc_1v_1 + cc_2v_2 + \ldots + cc_nv_n \in \mathrm{span}(S).$$

□

Thus, spans are indeed subspaces. The reason that we say a set S generates the span of S is that it turns out that the span of S is the smallest subspace of V containing S.

Example. *The set of standard basis vectors* $e_1, \ldots e_n$ *spans* $\mathbb{R}^n$.

Example. *The set* $\{1, x, x^2 \ldots, x^n, \ldots\}$ *generates* $\mathcal{P}(F)$ *for any field* F.

Example. *Given two vectors in* $\mathbb{R}^3$*, their span is either the set* $\{0\}$ *(if both vectors are* 0*), a line passing through one of them and the origin (if both of them lie on this line), or a plane (if they do not lie on a common line through the origin).*

Example. *The vectors* $(1, 2)$ *and* $(3, 4)$ *generate* $\mathbb{R}^2$*. Indeed, if* $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2$*, then we saw that for a square matrix* A*, the system* $Ax = b$ *is solvable if and only if* b *is in the span of the columns of* A*, and this is only true for all vectors* b *if the matrix is invertible (we argued this in the big proof in Section 2 of Lecture 7), so that this is equivalent to*

$$\det \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = -2 \neq 0,$$

which is true.

Example. *Continuing the last example, a set of* n *vectors in* $\mathbb{R}^n$ *spans all of* $\mathbb{R}^n$ *if and only if the determinant of the matrix with these vectors as its columns is non-zero. For example,*

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 0 & 9 \end{pmatrix} = -48 \neq 0,$$

so the vectors $(1, 4, 7)$*,* $(2, 5, 0)$*, and* $(3, 6, 9)$ *span* $\mathbb{R}^3$.

However,

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 0,$$

so $(1, 4, 7)$*,* $(2, 5, 8)$*, and* $(3, 6, 9)$ *do* **not** *span* $\mathbb{R}^3$.

Another way to see that this is true is to note that $-(1, 4, 7) + 2(2, 5, 8) = (3, 6, 9)$*, so that* $(3, 6, 9) \in \operatorname{span}((1, 4, 7), (2, 5, 8))$*. Thus, "adding in"* $(3, 6, 9)$ *doesn't make the span any larger.*

Example. *The polynomial* $2x^3 - 2x^2 + 12x - 6$ *is in the span of* $x^3 - 2x^2 - 5x - 3$ *and* $3x^3 - 5x^2 - 4x - 9$ *in* $\mathcal{P}(\mathbb{R})$*. Indeed, this is the same as saying that there are* a, b *with* $a(x^3 - 2x^2 - 5x - 3) + b(3x^3 - 5x^2 - 4x - 9) = (a + 3b)x^3 + (-2a - 5b)x^2 + (-5a - 4b)x +$

$(-3a-9b) = 2x^3 - 2x^2 + 12x - 6$, *or*

$$\begin{cases} a+3b=2 \\ -2a-5b=-2 \\ -5a-4b=12 \\ -3a-9b=-6 \end{cases}.$$

Using Gaussian elimination, we find that this is a consistent system of equations with solution $a=-4$, $b=2$.

Example. *We determine whether* $(-2,0,3)$ *is a linear combination of* $(1,3,0)$ *and* $(2,4,-1)$. *This is equivalent to the system* $Ax=b$ *with* $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 0 & -1 \end{pmatrix}$ *and* $b = \begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}$ *being consistent. Row reducing shows that this is indeed the case, and moreover explicitly finds the linear combination*

$$4(1,3,0) - 3(2,4,-1) = (-2,0,3).$$