LECTURE 5

Finitely Generated Vector Spaces

We are now in position to prove some general theorems about finite dimensional vector space that will be crucial to a number of applications.

But before starting on this, let me try to explain again, in a different way, our approach. The one habit I've been trying to wean you of is the an over-reliance upon concrete examples to develop your understanding. The vector space \mathbb{R}^n is a very concrete and familar example of a vector space over a field. To do calculations in this setting all you need to do is apply arithmetic (over and over and over). On the other hand, there are a number of other sets can be endowed with operations of scalar multiplication and vector addition so that they behave like \mathbb{R}^n . So we have a certain dichotomy here; a concrete and familar object, \mathbb{R}^n , and an associated set of patterns (the axioms of a vector space). What we are trying to do is deduce things from the patterns (axomatic vector space structure) that must be true for any object that satisfies the basic set of patterns. This allows us to say a whole lot about a whole lot of situations fairly succinctly.

However, in applications, one generally works in one particular situation at a time. Sometimes, for example, in freshman physics, one represents the points in space as list of three integers so often that people forget all the underlying apparatus that goes into interpreting a list three numbers as a point in space. But, in fact, one is using the calculational setting of \mathbb{R}^3 to deduce things about vector-like objects in space.

The trouble though with a purely abstract point of view is there is no means to calculate things. Think for a minute one how you would describe the location of particular objects in space without first setting up a coordinate system. Well, it works like this

points in space + coordinate system \Rightarrow a representation of point in space as elements of \mathbb{R}^3

Once we represent points in space by elements of \mathbb{R}^3 we can start to do cacluations.

Last time we began setting up the rudiments of a more general procedure

elements of a vector space V over a field \mathbb{F} + a basis for V \Rightarrow a representation of $v \in V$ as an element in \mathbb{F}^n

Once we specify the rules for arithmetic in the field \mathbb{F}^n , such a "coordinatization" of a vector space V over \mathbb{F} will allow us to calculate things in V. Indeed, the most important reason for introducing the notion of a basis is that the notion is essential to making questions about abstract vector spaces calculable.

Okay, here is a simple but useful lemma.

LEMMA 5.1. If $\{v_1, \ldots, v_m\}$ is a linearly dependent set and if $\{v_1, \ldots, v_{m-1}\}$ is a linearly independent set then v_m can be expressed as a linear combination of v_1, \ldots, v_{m-1} .

Proof. By hypothesis, there is a dependence relation of the form

(*) $\lambda_1 v_1 + \dots + \lambda_{m-1} v_{m-1} + \lambda_m v_m = \mathbf{0}_V$

with at least one of the coefficients $\lambda_i \neq 0_{\mathbb{F}}$. If it happened that $\lambda_m = 0_{\mathbb{F}}$, then we have

$$\lambda_1 v_1 + \dots + \lambda_{m-1} v_{m-1} = \mathbf{0}_V$$

with at least one of the λ_i , $1 \leq i \leq m-1$ not equal to $0_{\mathbb{F}}$. But then (**) would provide a dependence relation amongst the v_1, \ldots, v_{m-1} . But this is forbidden, since by hypothesis, the vectors v_1, \ldots, v_{m-1} are linear independent. Thus, we cannot have both (*) and $\lambda_m = 0_{\mathbb{F}}$. Since $\lambda_m \neq 0_{\mathbb{F}}$, we can divide each term in (*) by λ_m , to get

$$\frac{\lambda_1}{\lambda_m}v_1 + \dots + \frac{\lambda_{m-1}}{\lambda_m}v_{m-1} + v_m = \mathbf{0}_V$$

or

$$v_m = -\frac{\lambda_1}{\lambda_m}v_1 - \dots - \frac{\lambda_{m-1}}{\lambda_m}v_{m-1}$$

Thus, v_m is a linear combination of v_1, \ldots, v_{m-1} .

THEOREM 5.2. Every finitely generated vector space has a basis.

Proof. Let V be a vector space generated by n non-zero vectors $v_1 \ldots, v_m$. By Theorem 3.3, any set of m + 1 vectors in V must be linearly dependent. On the other hand, the set $\{v_1\}$ is certainly independent. We will now systematically generate a basis for V.

- Consider $\{v_1\}$. Since $v_1 \neq 0$ this is a linearly independent set.
- Consider $\{v_1, \ldots, v_m\}$. If this set is linearly independent, then will constitute a basis for V, since we already presume that it generates V.

If this set is not linearly independent, then there is a dependence relation

$$\beta_1 v_1 + \dots + \beta_m v_m = \mathbf{0}_V$$

with at least one non-zero coefficient. By reordering the vectors, if necessary, we can presume that $\beta_m \neq 0_{\mathbb{F}}$. But then we can express

$$v_m = -\frac{\beta_1}{\beta_m} v_1 - \dots - \frac{\beta_{m-1}}{\beta_m} v_{m-1}$$

And this will allow us to express every vector $v \in V = span_{\mathbb{F}}(v_1, \ldots, v_m)$ as a linear combination of v_1, \ldots, v_{m-1} .

So either $\{v_1, \ldots, v_m\}$ is a basis or we can re-express V as $span_{\mathbb{F}}(v_1, \ldots, v_{m-1})$.

- We can repeat the logic of the preceding step. Either $\{v_1, \ldots, v_{m-1}\}$ is a basis, or we can re-express V (after a suitable reordering of the vectors) as $span_{\mathbb{F}}(v_1, \ldots, v_{m-2})$.
- Repeating this process we will eventually either reach $V = span_{\mathbb{F}}(v_1)$ from which we'll conclude that $\{v_1\}$ is a basis for V or somewhere along the way we reached a generating set $\{v_1, \ldots, v_j\}$ consisting of linearly independent vectors. In the latter case, we can take $\{v_1, \ldots, v_j\}$ as a basis for V.

THEOREM 5.3. Let $V = span_{\mathbb{F}}(v_1, \ldots, v_m)$ be a finitely generated vector space. Then a basis for V can be selected from among the set of generators $\{v_1, \ldots, v_m\}$. In other words, any set of generators for a finitely generated vector space V contains a basis for V.

Proof. The construction of such a basis is given in the proof of Theorem 5.2.

THEOREM 5.4. Let $\{v_1, \ldots, v_k\}$ be a linearly independent set of vectors in a finitely generated vector space V. If v_1, \ldots, v_k is not a basis of V then there exist other vectors v_{k+1}, \ldots, v_m , such that $\{v_1, \ldots, v_m\}$ is a basis for V.

Proof. By Theorem 7.2, V, being a finitely generated vector space, has a basis, say $\{b_1, \ldots, b_n\}$. From Theorem 3.3 we know that the cardinality of any set of linearly independent vectors in V can not exceed the number of generators of V. Since $V = span_{\mathbb{F}}(b_1, \ldots, b_n)$ we must have $k \leq n$.

2

Now if. in fact, k = n, each $\{v_1, \ldots, v_n, b_i\}$ must be a linearly dependent set. And so by Lemma 5.1, we can express each b_i as a linear combination of v_1, \ldots, v_n . So $\{v_1, \ldots, v_n\}$ is, in fact, a linearly independent set that generates V; hence it is a basis.

Next suppose k = n-1. Then some $b_i \notin span_{\mathbb{F}}(v_1, \ldots, v_{n-1})$ (otherwise, all b_i would lie in $span(v_1, \ldots, v_{n-1})$ and so $\{v_1, \ldots, v_{n-1}\}$ would be a linear independent set of (n-1) vectors generating the vector space $V = span_{\mathbb{F}}(b_1, \ldots, b_n)$ of dimension n, which is impossible). But then $\{v_1, \ldots, v_{n-1}, b_i\}$ will be a linearly independent set generating V, and hence will be a basis for V.

We now suppose that n-k > 1, and the induction hypothesis that the theorem holds whenever the difference between dim V and the number of vectors $\{v_1, \ldots, v_k\}$ is less that n-k. Since $span(v_1, \ldots, v_k) \neq V$, there has to be some basis vector $b_i \notin span(v_1, \ldots, v_k)$. Then, as before $\{v_1, \ldots, v_k, b_i\}$ is a linearly independent set of k + 1 vectors. By the inductive hypothesis any linearly independent set of k + 1 vectors can be extended to a basis for V; for

$$n - (k + 1) = n - k - 1 < n - k$$

We can thus complete the set $\{v_1, \ldots, v_k, b_i\}$ to a basis $\{v_1, \ldots, v_k, b_i, b_{k+i+1}, \ldots, b_n\}$ of V. This then completes $\{v_1, \ldots, v_k\}$ to a basis of V.

Given two subsets S and T of a vector space V, we construct two additional subsets of V; viz,

$$S \cap T = \{ v \in V \mid v \in S \text{ and } v \in T \}$$

$$S \cup T = \{ v \in V \mid v \in S \text{ or } v \in T \}$$

When S and T are, in fact, subspaces of V, then natural questions to pose would be Is $S \cap T$ a subspace of V? and Is $S \cup T$ a subspace of V? These two questions were already posed in problems 4 and 5 of Homework Set 1. The answers are easy enough to answer directly here.

LEMMA 5.5. Let S and T be two subspaces of a vector space V over a field \mathbb{F} . Then $S \cap T$ is a subspace of V.

Proof.

It suffices to show that every linear combination of two elements of $S \cap T$ is an element of $S \cap T$. Suppose then that $u, v \in S \cap T$ and $\alpha, \beta \in \mathbb{F}$. Then

> $\alpha u + \beta v \in S$ since in particular $u, v \in S$ and S is a subspace $\alpha u + \beta v \in T$ since in particular $u, v \in T$ and T is a subspace

So $\alpha u + \beta v \in S \cap T$ and the proposition follows.

LEMMA 5.6. Suppose S and T are subpaces of a vector space V over a field \mathbb{F} . Then $S \cup T$ need not be a subspace.

Proof. It suffices to find one counter-example. Consider the following two subsets of \mathbb{R}^2

$$\begin{array}{rcl} S & = & \{ [x,0] \mid x \in \mathbb{R} \} & , \\ T & = & \{ [0,y] \mid y \in \mathbb{R} \} & . \end{array}$$

The vectors $u = [1, 0] \in S$ and $v = [0, 1] \in T$, certainly lie in $S \cup T$. However, their vector sum

$$u + v = [1, 1]$$

does not lie in $S \cup T$, because every element of $S \cup T$ has to be a vector that is either of the form [x, 0] or the form [0, y].

There is, however, another way to get a subspace from two subspaces

DEFINITION 5.7. Let S and T be two subspaces of a vector space V. Set

$$S + T = \{v = s + t \mid s \in S \text{ and } t \in T\}$$

What we've done here, of course, is defined our way around the problem underlying the counter-example in Lemma 5.6. That is to say, we have defined S + T in such a way as to guarantee that every linear combination of elements of S + T is another element of S + T. Indeed if $v_1 = s_1 + t_1$ and $v_2 = s_2 + t_2$ are arbitrary elements of S + T, then

$$\alpha v_1 + \beta v_2 = \alpha \left(s_1 + t_1 \right) + \beta \left(s_2 + t_2 \right) = \left(\alpha s_1 + \beta s_2 \right) + \left(\alpha t_1 + \beta t_2 \right) \in S + T$$

since $\alpha s_1 + \beta s_2 \in S$ and $\alpha t_1 + \beta t_2 \in T$.

Recall we defined subspaces precisely so that we could focus on subsets of a vector space V that were also vector spaces. The following lemma has similar utility, if you start in the category of finitely generated vector spaces and begin looking at subspaces, then you stay inside the category of finitely generated vector spaces.

LEMMA 5.8. If S is a nontrivial subspace of a finitely generated vector space V, then S itself is finitely generated.

Proof. Since S is nontrivial, it has at least one non-zero vector, say v_1 . Since S is a subspace, the vector generates a subspace $span_{\mathbb{F}}(v_1)$ of S. If $span_{\mathbb{F}}(v_1) = S$, then S is generated by v_1 and we are done.

If $S \neq span_{\mathbb{F}}(v_1)$, then there must be a non-zero vector $v_2 \in S$ that is not in $span_{\mathbb{F}}(v_1)$. In fact, $\{v_1, v_2\}$ must be linearly independent in this situation. For if

$$\alpha_1 v_1 + \alpha_2 v_2 = \mathbf{0}_V \qquad \alpha_1 \text{ and } a_2 \text{ not both } \mathbf{0}_{\mathbb{F}}$$

we would have

$$\alpha_1 v_1 = -\alpha_2 v_2$$

Which would either put $v_2 \in span_{\mathbb{F}}(v_1)$ if $\alpha_2 \neq 0$ (a contraction), or would put $\alpha_1 = 0_{\mathbb{F}}$ if $\alpha_2 = 0_{\mathbb{F}}$ (since v_1 is by assumption non-zero). If $span_{\mathbb{F}}(v_1, v_2) = S$, then S is generated by two vectors and we are done.

Otherwise, there has to be a third non-zero vector $v_3 \notin span_{\mathbb{F}}(v_1, v_2)$. By essentially the same argument as in the preceding paragraph, the vectors $\{v_1, v_2, v_3\}$ will have to be linearly independent in this case. And we can continue this process for finding more and more linearly independent vectors v_1, \ldots, v_i is S.

The point though is that this process has to terminate after finitely many steps. Because, by Theorem 5.2 a finitely generated vector space always has a basis. Say V has a basis $\{b_1, \ldots, b_n\}$. Then by Lemma 4.1, if i > n, then and set of i vectors in S (hence in V) must be linearly dependent. On the other hand, the only way the above algorithm can terminate is for $S = span(v_1, \ldots, v_i)$ for some i. We conclude that this indeed must be what happens for some $i \le n$, and so S is finitely generated.

THEOREM 5.9. Let S and T be finitely generated subspaces of a vector space V. Then $S \cap T$ and S + T are finitely generated subspaces, and we have

$$\dim (S+T) + \dim (S \cap T) = \dim S + \dim T$$

Proof. Let us first consider the case when $S \cap T$ is $\{\mathbf{0}_V\}$. Choose a basis $\{s_1, \ldots, s_m\}$ for S and a basis $\{t_1, \ldots, t_n\}$ for T. I claim $\{s_1, \ldots, s_m, t_1, \ldots, t_n\}$ is a basis for S + T. Indeed, vvery vector in $v \in S + T$ can then be expressed in the form s + t with $s \in S$ and $t \in T$. Since every $s \in S$ can be written as a unique linear combination of s_1, \ldots, s_m and every $t \in T$ can be expressed as a linear combination of t_1, \ldots, t_n , we'll have

$$v = s + t = (a_1s_1 + \dots + a_n) + (b_1t_1 + \dots + b_nt_n)$$

= $a_1s_1 + \dots + a_ms_m + b_1t_1 + \dots + b_nt_n$

so S + T is generated by $\{s_1, \ldots, s_m, t_1, \ldots, t_n\}$. I claim (under the hypothesis that $S \cap T = \{\mathbf{0}_V\}$) that $\{s_1, \ldots, s_m, t_1, \ldots, t_n\}$ is a linearly independent set. Suppose we had a dependence relation amongst these vectors

(*)
$$\alpha_1 s_1 + \dots + \alpha_m s_m + \beta_1 t_1 + \dots + \beta_n t_n = \mathbf{0}_V$$
 (with at least one coefficient nonzero)

Then, moving the terms involving the t_i 's to the right hand side we'll have

$$\alpha_1 s_1 + \dots + \alpha_m s_m = -\beta_1 t_1 - \dots - \beta_n t_n$$

Now the left hand side is in S while the right hand side is in T. By hypothesis, only vector common to both S and T is $\mathbf{0}_V$. Thus,

$$\alpha_1 s_1 + \dots + \alpha_m s_m = \mathbf{0}_V = -\beta_1 t_1 - \dots - \beta_n t_n$$

Since the basis vectors $\{s_1, \ldots, s_n\}$ are linearly independent the equality on the left forces $\alpha_1 = 0_{\mathbb{F}}, \alpha_2 = 0_{\mathbb{F}}, \ldots, \alpha_m = 0_{\mathbb{F}}$; and similarly since the basis vectors $\{t_1, \ldots, t_n\}$ are linearly independent the equality on the right forces $\beta_1 = 0_{\mathbb{F}}, \beta_2 = 0_{\mathbb{F}}, \ldots, \beta_n = 0_{\mathbb{F}}$. But this contradicts the existence of the dependence relation (*). Thus, when $S \cap T = \{\mathbf{0}_V\}$, the vectors $\{s_1, \ldots, s_m, t_1, \ldots, t_n\}$ will form a basis for S + T. In this case, we have

$$\dim (S+T) + \dim (S \cap T) = \# \{s_1, \dots, s_m, t_1, \dots, t_n\} + 0 = m + n$$

while

$$\dim (S) + \dim (T) = \# \{s_1, \dots, s_m\} + \# \{t_1, \dots, t_n\} = m + n$$

and so the statement of the theorem is confirmed.

Let us now consider the case when $S \cap T \neq \{\mathbf{0}_V\}$. $S \cap T$ is subspace of a finitely generated subspace $(S \cap T \subset S \text{ which by hypothesis is finitely generated}), S \cap T$ is finitely generated and so has a basis $\{u_1, \ldots, u_k\}$, where $k = \dim (S \cap T)$. Since $\{u_1, \ldots, u_k\}$ is a set of linear independent vectors in S and so, by Theorem 5.4, it can can be extend to a basis for S. Similarly, regarding $\{u_1, \ldots, u_k\}$ as a set of linear independent vectors in T, it can be completed to a basis for T. So doing we set

$$\{u_1, \dots, u_k, v_1, \dots, v_m\} , \quad \text{a basis for } S$$
$$\{u_1, \dots, u_k, w_1, \dots, w_n\} \quad , \quad \text{basis for } T$$

Then certainly

$$v \in S+T \Rightarrow v = (\alpha_1 u_1 + \dots + \alpha_k u_k + \alpha_{k+1} v_{k+1} + \dots + \alpha_m v_m) + (\beta_1 u_1 + \dots + \beta_k u_k + \beta_{k+1} w_{k+1} + \dots + \beta_n w_n)$$

$$\Rightarrow v = (\alpha_1 + \beta_1) u_1 + \dots + (\alpha_k + \beta_k) u_k + \alpha_{k+1} v_{k+1} + \dots + \alpha_m v_m + \beta_{k+1} w_{k+1} + \dots + \beta_n w_n$$

$$\Rightarrow v \in span_{\mathbb{F}} (u_1, \dots, u_k, v_1, \dots, v_m, w_1, \dots, w_n)$$

And its just as easy to show that

$$v \in span_{\mathbb{F}}(u_1, \ldots, u_k, v_1, \ldots, v_m, w_1, \ldots, w_n) \quad \Rightarrow \quad v \in S + T$$
.

 \mathbf{So}

$$S + T = span_{\mathbb{F}}(u_1, \dots, u_k, v_1, \dots, v_m, w_1, \dots, w_n)$$

I claim $\{u_1, \ldots, u_k, v_1, \ldots, v_m, \ldots, w_1, \ldots, w_n\}$ not only generates S+T, it is, in fact, a basis for S+T. To show this, we just need to demonstrate the vectors $\{u_1, \ldots, u_k, v_{k+1}, \ldots, v_m, w_{k+1}, \ldots, w_n\}$ form a linearly independent set. Suppose

(*)
$$\alpha_1 u_1 + \dots + \alpha_k u_k + \beta_1 v_1 + \dots + \beta_m v_m + \gamma_1 w_1 + \dots + \gamma_n w_n = \mathbf{0}_V$$

Then we have

$$\beta_1 v_1 + \dots + \beta_m v_m = -\alpha_1 u_1 - \dots - \alpha_k u_k - \gamma_1 w_1 - \dots - \gamma_n w_n$$

Now the left hand side is manifestly in S while the right hand side is manifestly in T. So both sides must be in $S \cap T$. This being the case, we can express either side of as linear combination of only the vectors u_1, \ldots, u_k . But then we must have an identity like

$$\beta_1 v_1 + \dots + \beta_m v_m = \lambda_1 u_1 + \dots + \lambda_k u_k$$

or

$$\beta_1 v_1 + \dots + \beta_m v_m - \lambda_1 u_1 - \dots - \lambda_k u_k = \mathbf{0}_V$$

But the vectors $\{u_1, \ldots, u_k, v_1, \ldots, v_m\}$ are a bais for S and so are linearly independent. This circumstance forces all the coefficients $\beta_1, \ldots, \beta_m = 0_{\mathbb{F}}$. Similarly, we could rewrite (*) as

$$\gamma_1 w_1 + \dots + \gamma_n w_n = -\alpha_1 u_1 - \dots - \alpha_k u_k - \beta_1 v_1 - \dots - \beta_m v_m$$

and deduce from this that $\gamma_1, \ldots, \gamma_n$ must all be equal to $0_{\mathbb{F}}$. Then given that $0_{\mathbb{F}} = \beta_1 = \beta_2 = \cdots = \beta_m$ and $0_{\mathbb{F}} = \gamma_1 = \gamma_2 = \cdots = \gamma_n$, we can further conclude from (*) that $0_{\mathbb{F}} = \alpha_1 = \alpha_2 = \cdots = \alpha_k$, since the vectors v_1, \ldots, v_k are linearly independent. Thus, since each of the coefficients must be equal to $0_{\mathbb{F}}$, (*) can not be a bone fide dependence relation, and so the vectors $\{u_1, \ldots, u_k, v_1, \ldots, v_m, w_1, \ldots, w_n\}$ are linearly independent and so provide a basis for S + T. We then have

$$\dim (S+T) = k + m + n$$

where

$$k = \dim (S \cap T)$$

$$m = \dim (S) - \dim (S \cap T)$$

$$n = \dim (T) - \dim (S \cap T)$$

And so

$$\dim (S+T) = \dim (S \cap T) + \dim (S) - \dim (S \cap T) + \dim (T) - \dim (S \cap T)$$

from which the statement of the theorem follows.

1. Infinite Dimensional Vector Spaces

We have just seen that all finitely generated vectors spaces are finite-dimensional. In this course, we will not say much about vectors spaces that are not finitely generated. Yet, such vector spaces are extremely important to mathematics and its applications. So I'll take a minute or so here to talk about how the notion of bases and dimension is dealt with in the case when a vector space V is not finitely generated.

To get this discussion moving forward we need to first adopt a standard axiom of set theory: the Axiom of Choice.

AXIOM 5.10. Given any family \mathcal{F} of mutually disjoint nonempty sets, there exists at least one set that contains exactly one element of each set in \mathcal{F} .

This axiom turns out to be equivalent to the following property of partially ordered sets.

LEMMA 5.11 (Zorn). If S is any non-empty partially ordered set in which every totally ordered subset in has an upper bound, then S has a maximal element.

THEOREM 5.12. Every non-trivial vector space V has a basis.

Proof. Let L be a set of linearly independent vectors in V (e.g. $L = \{v_1\}$ for some non-trivial vector in V). Let \mathcal{J} be the set of linearly independent subsets of V that contain L, partially ordered by inclusion. Then \mathcal{J} is non-empty because $L \in \mathcal{J}$. Moreover, if $J_1 \subset J_2 \subset J_3 \subset \cdots$ is a totally ordered chain of subsets in \mathcal{J} , then

$$M = \bigcup J_i$$

is also a linearly independent subset of V containing L, for if

$$v_1v_1 + \dots + r_nv_n = 0$$

with $v_i \in J_{\alpha(i)}$, Let $\alpha = \max \{\alpha(i)\}$ (there are at most *n* distinct *a*(*i*), and each is in \mathbb{N} , so we can find α). But then all the vectors $\{v_1, \ldots, v_n\}$ are in J_{α} . Since J_{α} is a linearly independent set r_1, \ldots, r_n must all equal $0_{\mathbb{F}}$. Hence, no dependence relation can exist amongst the vectors in *M* so *M* is a linearly independent set containing *L*; hence *M* is in \mathcal{J} . It is also clear that by construction the *M* is a maximal element of for

the chain $J_1 \subset J_2 \subset J_3 \subset \cdots$. Since every chain in \mathcal{J} has a maximal element, \mathcal{J} itself has an maximal element, call it B.

Now set W = span(B). If W = V, then B is a linearly independent set spanning V and so B is a basis. Suppose to the contrary, $V \neq W$. Then there exists $v \in V - W$. But then we can have no dependence relation of the form

$$av + a_1v_1 + \dots + a_nv_n = 0$$
 with $v_1, \dots, v_n \in B$

So $\{v, v_1, \ldots, v_n\}$ is a linearly indepedent set containg L, contradicting the maximality of B, Thus, no such v can exist and so V = span(B).

REMARK 5.13. In fact, the statement that every vector space has a basis is equivalent to the Axiom of Choice. (We've already seen how the Axiom of Choice, via Zorn's Lemma, allows one to deduce the existence of a basis for a general vector space. On the other hand, it turns out that if one postulates that every vector space has a basis, then the Axiom of Choice can be deduced as a consequence.)