2-12-1994

Linear independence and Bases.

Definition. Let V be an F-vector space, and let $S \subset V$. S is **linearly independent** if $v_1, \ldots, v_n \in S$, $a_1, \ldots, a_n \in F$, and

$$a_1 v_1 + \cdots + a_n v_n = 0 \quad \text{implies} \quad a_1 = \cdots = a_n = 0.$$

An equation like the one above is called a **linear relationship** among the v_i; if at least one of the coefficients a_i is nonzero, it is a **nontrivial linear relationship**. Thus, a set of vectors is independent if there is no nontrivial linear relationship among finitely many of the vectors.

A set of vectors which is not linearly independent is **linearly dependent**. I'll usually say "independent" and "dependent" for short.

Example. Any set containing the zero vector is dependent. For if $0 \in S$, then $1 \cdot 0 = 0$ is a nontrivial linear relationship in S. $\square$

Example. Consider $\mathbb{R}^2$ as a vector space over $\mathbb{R}$ in the usual way.

The vectors $\langle 1, 0\rangle$ and $\langle 0, 1\rangle$ are independent in $\mathbb{R}^2$. Suppose

$$a\langle 1, 0\rangle + b\langle 0, 1\rangle = \langle 0, 0\rangle, \quad a, b \in \mathbb{R}.$$

Then $\langle a, b\rangle = \langle 0, 0\rangle$, so $a = b = 0$.

More generally, the vectors

$$\langle 1, 0, 0, \ldots, 0\rangle, \langle 0, 1, 0, \ldots, 0\rangle, \ldots, \langle 0, 0, 0, \ldots, 1\rangle$$

are independent in $\mathbb{R}^n$. This set is called the **standard (ordered) basis** for $\mathbb{R}^n$. $\square$

Example. The vectors

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\},$$

are independent in $\mathbb{R}^3$.

Suppose

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

This is equivalent to the following set of linear equations:

$$\begin{array}{ccccccc} a & + & b & & & = & 0, \\ a & & & + & c & = & 0, \\ & & b & + & c & = & 0. \end{array}$$

You can verify that the solution is $a = b = c = 0$. Hence, the vectors are independent. $\square$

Remark. The equations above can be written in matrix form:

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

Many questions in linear algebra — such as determining whether a set of vectors is independent — can be handled by solving systems of equations. Systems of linear equations can be solved by a matrix algorithm called **row reduction**. □

Example. $\mathbb{R}[x]$ is a vector space over the reals. The set $1, x, x^2, \ldots$ is independent.

For if

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0,$$

it follows that $a_i = 0$ for all i. □

Definition. If V is an F-vector space, and $v_1, \ldots, v_n \in V$, a vector $v \in V$ is a **linear combination** of the v_i if

$$v = a_1 v_1 + \cdots + a_n v_n \quad \text{for some} \quad a_1, \ldots, a_n \in F.$$

Lemma. Let V be an F-vector space, and let $S \subset V$. S is linearly independent if and only if no $v \in S$ can be expressed as a linear combination of *other* vectors in S.

("Other" means vectors *other than* v itself.)

Proof. Suppose S is linearly independent. Suppose

$$v = a_1 v_1 + \cdots + a_n v_n$$

for $v_1, \ldots, v_n \in S$, $v_i \neq v$ for all i, $a_1, \ldots, a_n \in F$.

Then

$$0 = a_1 v_1 + \cdots + a_n v_n - 1 \cdot v$$

is a nontrivial linear relation among elements of S. This contradicts linear independence. Hence, v cannot be a linear combination of other vectors in S.

Conversely, suppose no $v \in S$ can be expressed as a linear combination of other vectors in S. Suppose

$$a_1 v_1 + \cdots + a_n v_n = 0,$$

where $a_i \in F$ and $v_i \in S$ for all i. I want to show that $a_i = 0$ for al i.

Suppose that at least one a_i is nonzero. Assume without loss of generality that $a_1 \neq 0$. Write

$$v_1 = -\frac{1}{a_1}(a_2 v_2 + \cdots + a_n v_n).$$

I've expressed v as a linear combination of other vectors in S $\times$. Hence, $a_i = 0$ for all i, and S is independent. □

Lemma. The intersection of subspaces is a subspace.

Proof. The intersection of a family of subgroups of a group is a subgroup, so the intersection of subspaces of an F-vector space V is a subgroup of $(V, +)$.

On the other hand, let $\{W_i\}$ be a family of subspaces of V, and let $W = \cap_i W_i$. If $a \in F$ and $w \in W$, then $w \in W_i$ for all i implies $aw \in W_i$ for all i. Hence, $aw \in W$, and W is closed under scalar multiplication. Therefore, W is a subspace. □

Definition. Let V be an F-vector space, and let $S \subset V$. The **span** of S (or the **subspace generated by** S) is the intersection of all subspaces which contain S. The span of S is denoted $\langle S \rangle$.

Since $S \subset V$, the intersection is nontrivial. By the previous lemma, the intersection is a subspace. It is clearly the smallest subspace containing S, in the sense that it is contained in any subspace which contains S.

If W is a subspace and $S \subset W$, S **spans** W if $W = \langle S \rangle$.

Lemma. Let V be an F-vector space, and let $S \subset V$.

$$\langle S \rangle = \left\{ \sum_{i=0}^{n} a_i v_i \mid n \geq 0, a_i \in F, v_i \in S \right\}.$$

That is, $\langle S \rangle$ consists of all linear combinations of elements of S.

Proof. Let

$$W = \{\sum_{i=0}^{n} a_i v_i \mid n \geq 0, a_i \in F, v_i \in S\}.$$

Since $\langle S \rangle$ is a subspace containing S, it must contain all linear combinations of elements of S. Thus, $W \subset \langle S \rangle$.

Conversely, observe that $W \neq \emptyset$, since $0 \in W$ (take the trivial linear combination of elements of S). Suppose $\{v_i\}_{i \in I}$, $\{v_j\}_{j \in J}$ are finite subset of S, and $a_i, b_j \in F$. The difference

$$\sum_{i \in I} a_i v_i - \sum_{j \in J} b_j v_j$$

is a linear combination of elements of S, so it's an element of W. Thus, W is a subgroup of $(V, +)$.

If $b \in F$ and $\sum_{i=0}^{n} a_i v_i \in W$, then

$$b \sum_{i=0}^{n} a_i v_i = \sum_{i=0}^{n} (b a_i) v_i \in W.$$

Therefore, W is closed under scalar multiplication.

It follows that W is a subspace. It obviously contains S, so $\langle S \rangle \subset W$. Therefore, $W = \langle S \rangle$. $\square$

Corollary. A set of vectors S spans a subspace W if every element of W is a linear combination of elements of S. $\square$

Example. Let

$$S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\} \subset \mathbb{R}^3.$$

Then

$$\langle S \rangle = \left\{ \begin{bmatrix} a \\ a+b \\ b \end{bmatrix} \,\middle|\, a, b \in \mathbb{R} \right\}. \square$$

Example. The span of $\{1, x, x^2\}$ in $\mathbb{Z}_3[x]$ consists of the 9 quadratic polynomials $ax^2 + bx + c$, where $a, b, c \in \mathbb{Z}_3$. $\square$

Example. The span of the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -4 \end{bmatrix} \right\}$$

is $\mathbb{R}^3$. In other words, if $\langle x, y, z\rangle$ is any vecotr in $\mathbb{R}^3$, there are real numbers a, b, c such that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = a \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + b \cdot \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + c \cdot \begin{bmatrix} 1 \\ 1 \\ -4 \end{bmatrix}.$$

In words, any vector is a linear combination of the three vectors in the set.

Obviously, there are sets of three vectors in $\mathbb{R}^3$ which don't span. (For example, take three vectors which are multiples of one another.) Geometrically, the span of a set of vectors in $\mathbb{R}^n$ can be a line through the origin, a plane through the origin, and so on. (The span must contain the origin, because a subspace must contain the zero vector.) $\square$

Linear independence has to do with the "nonredundancy" of a set of vectors. On the other hand, **spanning** has to do with being able to build everything in a subspace from a set of vectors. Here's how these ideas are related.

Definition. Let V be an F-vector space. A subset B of V is a **basis** if it is a maximal linearly independent set in V. "Maximal" means that if B' is another linearly independent set in V and $\mathcal{B} \subset \mathcal{B}'$, then $\mathcal{B} = \mathcal{B}'$.

Example. The vectors

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

are a basis for $\mathbb{R}^3$ over $\mathbb{R}$. So are

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}. \quad \square$$

Lemma. A basis for V spans V.

Proof. Let $\{v_i\}$ be a basis for V. Let $v \in V$. The set $\{v_i\} \cup v$ is dependent, by maximality of the basis. Hence, there exist $v_1, \ldots, v_n \in \{v_i\}$ and $a_1, \ldots, a_n, a \in F$, not all 0, such that

$$a_1 v_1 + \cdots + a_n v_n + av = 0.$$

If $a = 0$, then the equation is

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Since at least one of $a_1, \ldots, a_n, a$ is nonzero, this is a nontrivial linear relationship among elements of $\{v_i\}$, contradicting independence.

It follows that $a \neq 0$, and hence

$$v = -\frac{1}{a}(a_1 v_1 + \cdots + a_n v_n).$$

Therefore, every element of V is a linear combination of elements of $\{v_i\}$. $\square$

Bases are important objects in the study of vector spaces, because you can use them to represent vectors in different ways. Next, I will show that every vector space has a basis. Since a basis is a maximal set of some kind, it should not be surprising that the proof uses Zorn's lemma.

Theorem. Let V be an F-vector space, and let $S \subset V$, where $S \neq \emptyset$ and $S \neq 0$. Then S contains a maximal linearly independent set $\mathcal{B}$.

"Maximal" means that there is no independent subset of S which properly contains $\mathcal{B}$.

Proof. Let $\mathcal{S}$ be the set of linearly independent subsets of S. Since S is nonempty and nonzero, S contains a nonzero vector v. The set $\{v\}$ is a linearly independent subset of S, so $\mathcal{S} \neq \emptyset$.

Partially order $\mathcal{S}$ by inclusion. I'll use Zorn's lemma to show that $\mathcal{S}$ has maximal elements.

Let $\mathcal{T}$ be a chain in $\mathcal{S}$. Thus, elements of $\mathcal{T}$ are linearly independent subsets of S, and any two elements of $\mathcal{T}$ are comparable — one is contained in the other. Let

$$M = \bigcup_{T \in \mathcal{T}} T.$$

I claim M is an upper bound for $\mathcal{T}$ in $\mathcal{S}$.

If $M \in \mathcal{S}$, it is clearly an upper bound for the chain $\mathcal{T}$. So I only need to verify that M is in $\mathcal{S}$ — that is, M is a linearly independent subset of S.

Let $v_1, \ldots, v_n \in M$, and let $a_1, \ldots, a_n \in F$. Suppose

$$a_1 v_1 + \cdots a_n v_n = 0.$$

For each i, there is a $T_i \in \mathcal{T}$ such that $v_i \in T_i$. Since $\mathcal{T}$ is a chain, one of the T_i's contains all the others — without loss of generality, suppose that it is T_1. $a_1 v_1 + \cdots a_n v_n = 0$ is an equation involving elements of T_1, which is independent, so $a_i = 0$ for all i.

It follows that M is independent, so every chain in $\mathcal{S}$ has an upper bound. By Zorn's lemma, $\mathcal{S}$ has maximal elements. Such a maximal element is a maximal linearly independent subset of S. □

Remark. A trivial variation on this proof shows the following:

> If S is a subset of a vector space V and $T \subset S$ is a linearly independent, then there is a maximal linearly independent set M such that $T \subset M \subset S$.

In the proof above, simply let $\mathcal{S}$ be the collection of linearly independent subsets of S which contain T. Everything goes through exactly as above. □

Corollary. Every vector space has a basis.

(In order to avoid having to make a special case out of the zero vector space, people usually say that $V = \{0\}$ has the empty set as its basis, by convention.)

Proof. Apply the theorem with $S = V$. □

Corollary. A spanning set contains a basis.

More precisely, let $\mathcal{B}$ span a subspace W of a vector space V. Then some subset $\mathcal{B}' \subset \mathcal{B}$ is a basis for W.

Proof. Apply the theorem with $S = \mathcal{B}$. □

Remark. The proof of the last corollary shows that if the spanning set $\mathcal{B}$ contains n elements, then a basis for $\langle \mathcal{B} \rangle$ contains $\leq n$ elements. □

Corollary. A linearly independent subset of a vector space can be enlarged to a basis.

In other words, if $\mathcal{C}$ is linearly independent in V, there is a basis $\mathcal{B}$ of V such that $\mathcal{C} \subset \mathcal{B}$.

Proof. This follows from the remark immediately following the theorem. □

Example. The set

$$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -5 \\ 1 \end{bmatrix} \right\}$$

is dependent.

One way to see this is to write the third vector as a linear combination of the first two:

$$2 \cdot \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + 3 \cdot \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \\ 1 \end{bmatrix}.$$

Let W be the subspace spanned by $\mathcal{C}$. The results say that there is a subset of $\mathcal{C}$ which is a basis for $\mathcal{C}$. Indeed, the set

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \right\}$$

is independent, and spans W, so it's a basis for W. □