Lists and linear independence

Definitions of linear independence that apply only to *sets* of vectors make it harder to state and/or prove some important linear-algebraic facts. In these notes we approach linear independence in a more flexible way that does not have these deficiencies.

In these notes, "vector space" means "*real* vector space," but everything we say applies equally well to vector spaces over any field. We will often refer to real numbers as *scalars*; if we were talking about vector spaces over a different field, "scalars" would mean elements of that field.

Throughout these notes, V denotes a fixed but arbitrary vector space, "vector" means "element of V", and 0_V denotes the zero vector. An un-subscripted "0" means the scalar 0. We use the notation $\mathbf{N}$ for the set of natural numbers (positive integers).

Definition 1 Let X be a set. We will call a finite sequence $x_1, \ldots, x_n$ of elements of X a *list* of elements of X, or a list in X, and call the positive integer n the *length* of this list.[1] Given a list $x_1, \ldots, x_n$, the object x_i $(1 \leq i \leq n)$ is called the i^{th} *term* of the list, **not** the i^{th} *element* of the list. (The term x_i is still called an element <u>of the set X</u>; we simply do not call it an element of the given <u>list</u>.) An *n-term list* is a list of length n.

A list is a more general object than a finite, ordered, nonempty set. In a list $x_1, \ldots, x_n$, the terms x_i need not be distinct (there can be "repeats"), something that the definition of *set* does not allow. This is why we don't call x_i the i^{th} *element* of the list; only *sets* have elements.

Definition 2 Let L be a list $v_1, \ldots, v_n$ of vectors.

(a) A *linear combination* of L is a vector $v \in V$ for which there exists an n-term list of scalars $c_1, \ldots, c_n$ such that $v = c_1 v_1 + \cdots + c_n v_n$. (The sum in the last equation is also denoted $\sum_{i=1}^{n} c_i v_i$.)

(b) The *span* of L is the set of linear combinations of L:

$$\begin{aligned} \operatorname{span}(L) &:= \{v \in V : v \text{ is a linear combination of } L\} \\ &= \{v \in V : \ v = \sum_{i=1}^{n} c_i v_i \ \text{ for some list of scalars } c_1, \ldots, c_n\}. \end{aligned}$$

If $W = \operatorname{span}(L)$, we also say that L *spans* W.

[1] We use notation such as "$x_1, \ldots, x_n$" and "$\{1, \ldots, n\}$" rather than the more commonly seen "$x_1, x_2, \ldots, x_n$" and "$\{1, 2, \ldots, n\}$," to avoid giving the impression that n must be at least 2.

(c) We call the list L *linearly independent* if the only list of scalars $c_1, \ldots, c_n$ for which $c_1 v_1 + \cdots + c_n v_n = 0_V$ is the *trivial* one, meaning the list of scalars in which every term is 0. Otherwise we call L *linearly dependent*.

Proposition 3 *(a) Let L be a list of vectors $v_1, \ldots, v_n$ in which not all terms are distinct; i.e. for which $v_j = v_k$ for some $j, k \in \{1, \ldots, n\}$ with $j \neq k$.* (Obviously this is possible only if $n \geq 2$.) *Then L is linearly dependent.*

(b) If a list L in V is linearly independent, then all the terms of L are distinct.

Proof: (a) Let $j, k \in \{1, \ldots, n\}$ be such that $j \neq k$ but $v_j = v_k$. Let $c_1, \ldots, c_n$ be the list of scalars in which $c_j = 1$, $c_k = -1$, and $c_i = 0$ for all i other than j, k. Then this list of scalars is not the trivial list, but $\sum_{i=1}^{n} c_i v_i = v_j - v_k = 0_V$. Hence L is linearly dependent.

(b) This follows immediately from (a). ■

Thus, a list of vectors does not even have a *chance* of being linearly independent unless all its terms are distinct. This is the reason for the distinctness requirement in the definition below (one of the usual definitions of linear independence/dependence of a *set* of vectors):

Definition 4 Let $S \subseteq V$. We say that S is *linearly independent* if every list of distinct elements of S is linearly independent. Otherwise we say that S is *linearly dependent*.

Example. As an example of the convenience afforded by defining linear independence of *lists* of vectors, rather than just *sets* of vectors, consider the following. Every $m \times n$ matrix has n columns. Or does it? What are the columns of the 4×3 matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 3 & 4 & 3 \\ 4 & 5 & 4 \end{pmatrix} \quad ?$$

Each column is a vector in $\mathbf{R}^4$. Since the first and third columns are identical, the *set* of columns consists of only two vectors. But, looking at A, we see a third column sitting there. *Implicitly, we are regarding the columns of A as a 3-term* list *of elements of $\mathbf{R}^4$, not just as* set *of elements of $\mathbf{R}^4$*. The answer to our "Or does it?" is yes, as we originally thought. Whether all the columns of an $m \times n$ matrix are distinct or not, an $m \times n$ matrix has exactly n columns, because (without having said so explicitly in the past), we have always treated the columns of an $m \times n$ matrix as the terms of an n-term of a *list* of elements of $\mathbf{R}^m$.

Note that the *set* of columns of A—a two-element set—is linearly independent. But the *list* of columns is not. The column space of A—the span of the columns, as a subspace of $\mathbf{R}^4$—has dimension 2. Observe that this would be true also if the third column were 1.0001 times the first column, instead of being exactly 1 times the first column. In that case, the set of columns would be a three-element set rather than a two-element set. But either way, the columns form a three-term *list*.

Another concept more general than *set* that allows duplication is *multi-set*. In a multi-set, each element is counted with a *multiplicity* that can be any positive integer. The columns of the matrix A above form a multi-set with two (distinct) elements, one of which has multiplicity 2 and the other of which has multiplicity 1. But this multi-set still carries less information than the *list* of columns, because a multi-set does not take *order* into account. Consider the matrix

$$B = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 3 \\ 3 & 3 & 4 \\ 4 & 4 & 5 \end{pmatrix}.$$

The matrices A and B have the same *set* of columns, and even have the same *multi*-set of columns, yet are different matrices because the columns appear in different orders. The two *lists* of columns are different.

Remark 5 Since (by Proposition 3), the terms of any linearly independent list of vectors are distinct, any such list $v_1, \ldots, v_m$ can be identified with an ordered *set* $\{v_1, \ldots, v_m\}$ We will make this identification implicitly.

An ordered *basis* is then seen to be a (very special) linearly independent list. But, as the comparison of the matrices A and B in the example above shows, ordered *bases* are not the only important linearly independent lists. The ordering of a list of linearly independent vectors can matter even when the list does not span V

As (part of) Remark 5 shows, the following definition of "ordered basis" is equivalent to the one we've been using:

Definition 6 Assume that V is finite-dimensional and that $\dim(V) > 0$. An *ordered basis* of V is a linearly independent list (in V) that spans V.

Using the implicit identification in Remark 5, an ordered basis *is* a special type of *basis*, as the terminology suggests.

Proposition 7 *Assume that V has finite, positive dimension n. Then:*

(a) No list in V with more *than n terms can be linearly independent.*

(b) No list in V with fewer *than n terms can span V.*

(c) If L is a list in V with exactly *n terms, then L is linearly independent if and only if L spans V.*

In class we proved an analogous proposition for sets with more than n, fewer than n, or exactly n elements. Note that Proposition 7 is *stronger* than this analogous proposition, because it applies whether or not the terms of the list are distinct (equivalently, whether the number of terms is the same as the number of *distinct* terms).

Proof: We will make use of the folliwing result we proved previously as a corollary of the rank-plus-nullity theorem: If W is a finite-dimensional vector space and $T : W \to V$ is linear, then (a′) if $\dim(W) > \dim(V)$, T cannot be one-to-one; (b′) if $\dim(W) < \dim(V)$, then T cannot be onto; and (c′) if $\dim(W) = \dim(V)$, then T is one-to-one if and only if T is onto.

Let $v_1, \ldots, v_m$ be a list L of vectors in V. Define $T : \mathbf{R}^m \to V$ by

$$T(\begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}) = \sum_{i=1}^{m} c_i v_i \ .$$

It is easily checked that T is linear. (Students: wording like that is telling you to check, not to assume the truth of what was just stated.)

By definition of span(L), the range of T is precisely span(L). Hence T is onto if and only if L spans V. Furthermore, since T is a linear transformation, T is one-to-one if and only if $N(T) = \{0_{\mathbf{R}^m}\}$, which is equivalent to the statement that the only list $c_1, \ldots, c_m$ of scalars for which $\sum_i c_i v_i = 0_V$ is the trivial list of m scalars, which in turn, is equivalent to the statement that L is linearly independent.

(a) Assume that $m > n$. Then by (a′), T is not one-to-one, so $\mathsf{N}(T) \neq \{0_{\mathbf{R}^m}\}$, so L is linearly dependent.

(b) Assume that $m < n$. Then by (b′), T is not onto. Hence L does not span V.

(c) Assume that $m = n$. Then, by (c′), T is one-to-one if and only if T is onto. But, as seen earlier, "T is one-to-one" is equivalent to "L is linearly independent", while "T is onto" is equivalent to "span(L) = V." Hence L is linearly independent if and only if L spans V. ■

Corollary 8 *Assume that V has finite, positive dimension n. If L is an n-term linearly independent list in V, then L is an ordered basis of V.*

Proof: This follows immediately from Proposition 7(c). ■

Application to Problem 4 on Exam 1. (I'll be referring to my solutions handout, so get that out to follow along.)

The statement of this problem was:

> Let $u, v,$ and w be distinct vectors in a vector space V. Show that if $\{u, v, w\}$ is a basis of V, then so is $\{u + v + w, v + w, w\}$

In my solutions handout, I gave a "proof" that, as I mentioned there, had a gap that (after challenging you to find the gap!) I fixed with an *ad hoc* argument. We can now fix the proof a better way.

We start the same way as in the solutions handout: "Assume that $\{u, v, w\}$ is a basis of V. Then $\dim(V) = 3$."

Next, let L be the three-term list "$u + v + w,\ v + w,\ w$." In the solutions handout, what the portion of the argument from "Let $a, b, c \in \mathbf{R}$" through "Hence S is linearly independent" actually shows is that the *list* L is linearly independent. Hence, by Corollary 8, L is an ordered basis of V. ■

We could actually have filled the gap without using beyond Proposition 3 of these notes. After showing that the list L is linearly independent, we could have argued as follows:

> By Proposition 3, the terms of L are distinct. Hence $S := \{u+v+w, v+w, w\}$ is a three-element linearly independent *set* in the three-dimensional vector space V, so (by the result from class referred to in the solutions handout), S is a basis of V. ■

Some additional comments

• The word "list" is my own, shorter term for "finite sequence"; it's not official terminology.

• Mathematically there is no difference between (i) an n-term sequence $x_1, \ldots, x_n$ of elements of a set X, and (ii) an ordered n-tuple $(x_1, \ldots, x_n)$ of elements of X, also known as an element of $X \times X \times \cdots \times X$ (n copies of X), the n-fold Cartesian product of X with itself. There is sometimes a bit of a difference in the way we *think* about a list vs. an ordered n-tuple. In a list, we generally are thinking, "Here's the first term, here's the second term, here's the third term," etc. In an ordered n-tuple $(x_1, \ldots, x_n)$, we tend to think of all the x_i as being present "at once" (like the coordinates of a point $(x, y, z) \in \mathbf{R}^3$), rather than as elements being selected in a particular time-order.

- If we want to give a particular list a short name, say L, the notation "$x_1, \ldots, x_n$" does not work well in sentence like "Let $L = x_1, \ldots, x_n$ be a list of elements of X;" it looks at first like we're setting L equal to x_1. Ideally one should "protect" the list with symbols that keep the list, as an object, separate from anythiing else. Parentheses would do this, as in "Let $L = (x_1, \ldots, x_n)$," but that notation can be misleading if the x_i are distinct and we want to think of ths list as an ordered *set*. One notation that can be used is $\{x_i\}_{i=1}^n$, but this is not perfect either, as it can be misinterpreted as the *set* $\{x_i : 1 \leq i \leq n\}$, in which duplicates would count as the same element rather than as different terms of a sequence. None of these choices is ideal, so I just used the long-form notation "$x_1, \ldots, x_n$" in these notes, and avoided writing anything like "$L = x_1, \ldots, x_n$".