Lists, linear combinations, and linear independence

version: 10/5/2024, 22:30 EDT

Definitions of linear combinations and linear dependence/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 these topics 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. A parenthetic item in blue is generally a *side comment*; it's not actually part of the definition, proposition, etc. that it's sitting inside.

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 (where $1 \le i \le n$) 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 of the set X; we just don't call it an element of the given list.) An n-*term list* is a list of length n.

A list is a more general object than a finite, ordered, nonempty set. (For what "finite, ordered set" means, see Remark 10, paragraph 2.) 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 of vectors $v_1, \ldots, v_n$.

(a) A *linear combination* of the list L is a vector $v \in V$ for which there exists an n-term list of scalars $c_1, \ldots, c_n$ such that $v = c_1v_1 + \cdots + c_nv_n$. (The sum in the last equation is also denoted $\sum_{i=1}^{n} c_iv_i$.)

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

$$\begin{aligned} \text{span}(L) &:= \{v \in V : v \text{ is a linear combination of } L\} && (1)\\ &= \{c_1v_1 + \cdots + c_nv_n \;:\; c_1, \ldots, c_n \in \mathbf{R}\} && (2)\\ &= \Big\{v \in V :\ v = \sum_{i=1}^{n} c_iv_i \ \text{ for some list of scalars } c_1, \ldots, c_n\Big\}. && (3)\end{aligned}$$

[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.

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

(*Note*: The right-hand sides on lines (1), (2), and (3) above are just different ways of writing the same thing. We can go back and forth among them. Which one we choose, at any given moment, can depend on what we're trying to show or explain.)

(c) We call the list L *linearly independent* if the only list of scalars $c_1, \ldots, c_n$ for which $c_1v_1 + \cdots + c_nv_n = 0_V$ is the *trivial* list of scalars, meaning that $c_i = 0$ for each $i \in \{1, \ldots, n\}$ (which is often written as "$c_1 = c_2 = \ldots c_n = 0$" if n is assumed to be at least 2, or as "$c_1 = \ldots c_n = 0$" without that assumption). 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 $\sum_{i=1}^{n} c_iv_i = v_j - v_k = 0_V$. 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*.

Note that the empty set $\emptyset$ meets the definition of "linearly independent subset of V", since there *are* no lists of distinct elements of S. (Each of these nonexistent lists meets any condition whatsoever!).

Parts (a) and (b) of Definition 2 also have analogs for subsets of V:

Definition 5 (a) A *linear combination* of elements of S is a vector $v \in V$ that is a linear combination of some list in S.

(b) (i) If $S \neq \emptyset$, the *span* of S is the set of linear combinations of elements of S:

$$\begin{aligned} \text{span}(S) &:= \{v \in V : v \text{ is a linear combination of some list of elements of } S\} \quad (4) \\ &= \{c_1 v_1 + \cdots + c_n v_n \ : \ n \geq 1, \ \ v_1, \ldots, v_n \in S, \ \text{and } c_1, \ldots, c_n \in \mathbf{R}\} \quad (5) \\ &= \Big\{v \in V : \ v = \sum_{i=1}^{n} c_i v_i \ \text{ for some } n \geq 1, \\ &\qquad\qquad \text{some list of vectors } v_1, \ldots, v_n \in S, \\ &\qquad\qquad \text{and some list of scalars } c_1, \ldots, c_n \Big\}. \quad (6) \end{aligned}$$

(ii) We define $\text{span}(\emptyset) = \{0_V\}$.

In both cases (i) and (ii), if $W = \text{span}(S)$, we also say that S *spans* W (or *generates* W).

Remark 6 (Linear combinations: sets vs. lists) Note that for a vector v to be a linear combination of a list $v_1, \ldots, v_n$, every term v_i of the list must appear on the RHS (right-hand side) of the equation

$$v = c_1 v_1 + \ldots c_n v_n \ , \quad (7)$$

multiplied by a scalar c_i that may or may not be 0. For v to be a linear combination of elements of a set S, there is *no* requirement that every element of S appear in the RHS of (7). Such a requirement would cause difficulties when the set S is infinite, since we can't list all the elements of an infinite set. (However, for a *finite* nonempty set $S = \{v_1, \ldots, v_n\}$, the definition of "linear combination of elements of S" is equivalent to one that is identical to Definition 2(a), modulo replacing the words "the list L" by "elements of S". *Exercise*: Understand *why* these two definitions of " linear combination", for a **finite** nonempty set S, are equivalent.)

Exercise 7 (a) Show that if any (i.e. at least one of) the terms of a list L in V is the zero vector, then L is linearly dependent.

(b) Show that if a subset S of V contains the zero vector, then S is linearly dependent.

Proposition 8 *Let L be a list of vectors $v_1, \ldots, v_n$ in V and let S_L be the set of terms of L.* (Thus $S_L = \{v_i : 1 \leq i \leq n\} = \big\{v \in V : v = v_i \text{ for some } i \in \{1, \ldots, n\}\big\}$.) *Then:*

(a) $\text{span}(S_L) = \text{span}(L)$.

(Said another way: the span of L [the right-hand side of the equation above] is the same as the span of the set of distinct terms of L [the left-hand side of the equation above]).

(b) If L is linearly independent, so is S_L.

Equivalently, if S_L is linearly dependent, then so is L.

(c) Assume the terms of L are distinct. Then L is linearly independent if and only if S_L is linearly independent.

If you have a good understanding of the definitions involved, part (a) of this proposition may seem obvious. Writing down a formal proof without saying "it's obvious"—which it may *not* be to everyone for whom the ideas are new—is rather tedious, and you might find it difficult to do. The proof below is one that you may find tedious even to *read*, but I have supplied it for the sake of completeness, and because some of the ideas are ones that you (the student) might have trouble stating with precision on your own (without relying on words that make sense to *you* but that haven't been given a mathematical definition).

Proof: (a) Let $m = |S_L|$ (the number of *distinct* terms of L); thus $m \leq n$. Since a list has at least one term (by definition), S_L is not empty, so $m \geq 1$ and $\text{span}(S_L)$ is exactly the set of linear combinations of elements of S_L. We may list the elements $w_1, \ldots, w_m$ be the elements of S_L as $v_{i_1}, \ldots, v_{i_m}$ for some $i_1, \ldots i_m \in \{1, \ldots, n\}$ with $1 = i_1 < i_2 < \cdots < i_m$ (with both i_2 and i_m present in this chain of inequalities only if $m \geq 3$, and with i_2 present only if $m \geq 2$).

Let $a_1, \ldots, a_n \in \mathbf{R}$. Then

$$a_1 v_1 + \cdots + a_n v_n = b_1 w_1 + \cdots + b_m w_m \ ,$$

where, for each $j \in \{1, \ldots, m\}$, the coefficient b_j is the sum of the coefficients a_i for which $v_i = w_j$. Hence $\text{span}(L) \subseteq \text{span}(S_L)$.

Now let $b'_1, \ldots, b'_m \in \mathbf{R}$, and for $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$ let

$$a'_i = \begin{cases} b'_j & \text{if } i = i_j \, , \\ 0 & \text{otherwise.} \end{cases}$$

(Thus the list of scalars a'_i has the property that $a'_i = 0$ for all i [if any] satisfying $i_m < i \leq n$, and, if $m \geq 2$, for all i satisfying $i_j < i < i_{j+1}$ for some $j \in \{1, \ldots, m-1\}$. E.g. if we have $n = 10$, $m = 4$, $i_2 = 4$, $i_3 = 5$, and $i_4 = 7$, then the list $a'_1, \ldots a'_{10}$ is $b'_1\,, 0\,, 0\,, b'_2\,, b'_3\,, 0, b'_4\,, 0\,, 0\,, 0$.) Then $b'_1 w_1 + \cdots + b'_1 w_m = a_{i_1} v_{i_1} + \cdots + a_{i_m} v_{i_m} = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$ (where "$a_2 v_2$" is present only if $m \geq 2$). Hence $\text{span}(S_L) \subseteq \text{span}(L)$, completing the proof that $\text{span}(S_L) = \text{span}(L)$.

(b) Assume that the list L is linearly independent. Then by Proposition 3(b), the terms of L are distinct, so $|S_L| = n$, and the n elements of S_L are precisely $v_1, \ldots, v_n$.

Suppose that $k \geq 1$, that $w_1, \ldots, w_k$ are k distinct elements of S_L, and that $b_1, \ldots, b_k \in \mathbf{R}$ are such that

$$b_1 w_1 + \cdots + b_k w_k = 0_V. \tag{8}$$

Since $w_1, \ldots, w_k$ are distinct, for each $j \in \{1, \ldots, k\}$ there is a unique $i_j \in \{1, \ldots, n\}$ such that $w_j = v_{i_j}$. Since the vectors w_j are distinct, so are the indices $i_1, \ldots, i_k$. Reordering the list $w_1, \ldots, w_k$ if necessary, we may assume that $i_1 < i_2 < \cdots < i_k$ (with i_2 and i_k both present in this chain of inequalities only if $k \geq 3$, and with i_2 present only if $k \geq 2$). For $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, k\}$ let

$$a_i = \begin{cases} b_j & \text{if } i = i_j\,, \\ 0 & \text{otherwise.} \end{cases}$$

Then $b_1 w_1 + \cdots + b_k w_k = a_{i_1} v_{i_1} + \cdots + a_{i_k} v_{i_k} = a_1 v_1 + a_2 v_2 + \cdots + a_m v_m$ (with "$a_2 v_2$" present only if $k \geq 2$). Hence the hypothesis (8) implies that $\sum_{i=1} a_i v_i = 0_V$, which implies that $a_i = 0$ for each i since (by hypothesis) L is linearly independent. But then $b_j = 0$ for each $j \in \{1, \ldots, k\}$ as well.

Hence the set S_L is linearly independent.

(c) By part (b), the "$\Longrightarrow$" direction of this "iff" statement holds (whether or not the terms of L are distinct), so we need only prove the "$\Longleftarrow$" implication.

Assume that S_L is linearly independent. Let $v_1, \ldots, v_n$ be the terms of L. Suppose $a_1, \ldots, a_n \in \mathbf{R}$ are such that

$$a_1 v_1 + \cdots + a_n v_n = 0_V\,. \tag{9}$$

By hypothesis the v_i are distinct, so $v_1, \ldots v_n$ is a list of distinct elements of the linearly independent set S_L. Hence each a_i is 0. Hence L is linearly independent. ■

Example 9 As an example of the convenience afforded by defining linear independence of *lists* of vectors, rather than just of *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 $\mathbf{R}^4$. (Once we start talking *linear transformations*, it will becomer more convenient to write elements of $\mathbf{R}^n$ as *column vectors*; equivalently, as $n \times 1$ matrices.) 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 we always treat the columns of an $m \times n$ matrix as the terms of an n-term *list* of elements of $\mathbf{R}^m$. [Example continues on next page.]

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*. [End of Example 9.]

Remark 10 (Multi-sets and ordered sets) (Optional reading) 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.

Another related concept is that of an *ordered* finite set. An ordered finite set is simply a finite set with the elements listed in a chosen order. For example, the six expressions

$$\{1,2,3\},\ \{1,3,2\},\ \{2,1,3\},\ \{2,3,1\},\ \{3,1,2\},\ \{3,2,1\}$$

all represent the same *set*, but six different *ordered* sets.

When we write down a (finite) set by literally listing all its elements, or by indexing them by (say) $\{1,\ldots n\}$, as in "$\{v_1,\ldots,v_n\}$", we are forced to choose an order in which to write the elements. Sometimes it's clear from context that we're regarding the collection of elements simply as a *set*, not an *ordered* set. For example, a class roster usually lists the students in alphabetical order, and while this is convenient for record-keeping, everyone understands that there's no such thing as the first student in the class, the second student, etc. Unfortunately, there is no standard notation that distinguishes an ordered finite set from an unordered set (i.e. simply a *set*). *By default*, the notation "$\{v_1,\ldots,v_n\}$" means the *un*ordered set. If the writer's intended meaning is *ordered* set, he/she has to so.

To illustrate the different concepts discussed above—*set*, *ordered set*, *list*, and *multi-set*—consider the matrices

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

These two matrices 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. The two *ordered sets* formed by the columns of A may be written as

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \right\} \quad \text{and} \quad \left\{ \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \right\},$$

both of which are ways of writing the *set* of columns of A.

Remark 11 By Proposition 3, the terms of any linearly independent list of vectors are distinct. Hence any such list $v_1, \ldots, v_m$ can be identified with the *ordered set* $\{v_1, \ldots, v_m\}$. (See the second paragraph of Remark 10 for the meaning of "ordered set".) We will make this identification implicitly, treating such a list as an ordered set.

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 11 shows, the following definition of "ordered basis" is equivalent to the one we've been using:

Definition 12 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 11, an ordered basis *is* a special type of *basis*, as the terminology suggests.

Proposition 13 *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.*

Many textbooks, e.g. [1], state and prove an analogous proposition for <u>sets</u> with more than n, fewer than n, or exactly n <u>elements</u>. Note that Proposition 13 is *stronger* than this analogous proposition, because it applies whether or not the terms of the list are distinct (equivalently, whether or not the number of terms is the same as the number of *distinct* terms).

Note: The original version of this handout was written for my Fall 2023 class after the "Rank-Plus-Nullity Theorem"—which [1] calls the "Dimension Theorem"—had been covered.

In the interests of efficiency at the time, I used that theorem in the proof below. For my 2024 class, a different set of notes has a proof of Proposition 13 that does not rely on the Rank-Plus-Nullity Theorem.

Proof: We will make use of the following 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 take the writer's word!)

By definition of $\text{span}(L)$, the range of T is precisely $\text{span}(L)$. 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 $\mathsf{N}(T) = \{0_{\mathbf{R}^m}\}$, 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 "$\text{span}(L) = V$." Hence L is linearly independent if and only if L spans V. ■

Corollary 14 *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 13(c). ■

<u>Some additional comments</u>

• **The word "list" (as used in these notes) is my own, shorter term for "finite sequence"; it's not official terminology.** One reason I use "list" instead of "finite sequence" is to make sure students won't think I might be talking about the *infinite* sequences and series they learned about in Calculus 2, which are irrelevant to the discussion in these notes.

• 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 anything 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 this list as an ordered *set*. Furthermore, in a first course on linear algebra, the notation "$(x_1, \ldots, x_n)$" could lead some students to assume the entries of the n-tuple are *numbers*, if that's the only type of n-tuples a student has seen before.) 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 chose just to use the long-form notation "$x_1, \ldots, x_n$" in these notes, and avoided writing anything like "$L = x_1, \ldots, x_n$".

References

[1] Friedberg, Insel, and Spence, *Linear Algebra*, 5th edition. Pearson Education, Inc., 2019