Linearly Independent Study

Na'im Kalantar

June 5, 2017

1 Vectors

The theorems of linear algebra all follow from specific rules and properties of certain sets. Therefore, linear algebra can be applied to anything that can be modeled as a set with this structure.

The structure of interest is called the **vector space**, with elements of a vector space called **vectors**. Vectors can add with other vectors of the same kind, but can't multiply with each other. They can multiply with regular numbers. Anything is a vector space if it follows these rules:

1. A vector space needs to have addition, with the following rules:

    (a) The sum of two vectors in the vector space is also in the vector space.

    (b) For any vectors of the same kind, it doesn't matter what order you add them in: $u + v = v + u$, and $(u + v) + w = u + (v + w)$.

    (c) There's a special vector that has an unusual property: Anything plus this vector is just itself. In normal numbers, this vector is called 'zero', and we'll write it 0_v, the zero vector.: $v + 0_v = v$

    (d) Every vector v has another vector, v^{-1} or 'v-inverse', where $v + v^{-1} = 0_v$.

2. A vector space needs to multiply with regular numbers[1], with the following rules:

    (a) Every product of vectors is also in the vector space.

    (b) It doesn't matter what order you multiply things: if a and b are numbers, and v is a vector: $a(bv) = (ab)v$

    (c) There's a special number so that this number times a vector is itself. This number is kind of like 1: $1v = v$

    (d) Addition and multiplication interact like they do for normal numbers: $a(u + v) = au + av$, and $(a + b)v = av + bv$

---

[1] 'Regular numbers' here really means a **field**, which, like a vector space, is a set of anything that follows special rules. The real numbers, $\mathbb{R}$ and the complex numbers, $\mathbb{C}$, are the most interesting fields.

2   Some examples

These definitions are are very useful because it turns out that there are a lot of things that fit them. This means that if we use the rules above to prove something, it applies to all of them. Here are some examples of sets that are vector spaces because they fit the rules:

1. It's possible to think of regular numbers as vectors, where the vector space is the group of all numbers. If you like, you can check that the regular numbers fit the rules.

2. Points in space are vectors. This is true if you are thinking about normal three-dimensional space, something flatter like points on a line or points in a flat plane, or some kind of space with more than three dimensions. Thinking about space as a vector space and using the rules of linear algebra has important applications in physics.

3. Lists of numbers of the same size fit the rules, with defined addition like this: $(a_1, a_2, a_3) + (b_1, b_2, b_3) = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$. So you can think of the lists of numbers as vectors, and the group of all lists of the same size is a vector space.

4. Certain kinds of functions also fit the rules. Try polynomials of a certain degree, or continuous functions[2] that are defined from 0 to 1.

2.1   Proving things using the rules

If a statement logically follows from the rules, it must be true for all vector spaces. Here's an example of one such statement.

Theorem. *For every vector v, $-1 * v$ is its inverse: $-1 * v + v = 0_v$*

Proof. $-1 * v + v = -1 * v + 1 * v$ by 2c. $-1 * v + 1 * v = (-1 + 1)(v)$ by 2d. $(-1 + 1)(v) = 0v$. $0v = (0 + 0)v$ by 2d. $0v = 0v + 0v$ by 2c. $0v + (0v)^{-1} = 0v + 0v + (0v)^{-1}$ by adding the inverse of $0v$ to each side. $0 = 0v$ by definition of inverse. From what's been said so far, $-1 * v + v = (-1 + 1)v = 0v = 0_v \implies -1 * v + v = 0_v$. $\square$

3   Subspaces

Sometimes, subsets of vectors spaces follow the vector space rules, and are themselves vector spaces. Vector spaces that are contained in another vector space are called **subspaces**. For example, take all lists of numbers of size three which have the same first and last number – That is, lists of numbers which have the form (a, b, a). This subset is a subspace of the vector space of all lists of size three. (Check yourself!) For another example, take the set of lists of

---

[2]Functions that don't have jumps and don't go off to infinity

numbers of any size whose elements add up to zero – Lists that look like (a, b, c) where $a + b + c = 0$.

Obviously not all subsets of a vector space are subspaces. For an obvious example. The subset of polynomials with highest power x^2 that only contains $x^2 + 4x + 4$ and $2x^2 + 42x - \pi$ doesn't follow the rules: Addition of the two elements of the subset gives you something that isn't in the subset, and that breaks the first rule of vector space addition.

There are simple conditions on a subset of a vector space that guarantee it will be a subspace. They are that the subspace contains 0_v, that the subspace contains any sum of two elements in the subspace, and that the subspace contains the product of any number with any vector. (If you want to check, go through and check all the rules).

4 Span and Linear Independence

4.1 Span

All of the vector spaces that are described above and many other interesting ones have an infinite number of vectors in them. However, some vector spaces are subspaces of other larger spaces, and it feels like the subspaces should be 'smaller'. To make it easier to think about different vector spaces, their sizes, and their relationship to each other, we describe a space as the **span** of a set of vectors. When we take a set of vectors we always need them to be vectors in the same vector space – vectors that are of the same kind.

Take a set of vectors S. The span of S is another set. A vector is in this new set if, and only if, it can be written as the sum of multiples of elements of S. If $S = \{v_1, v_2, v_3\}$, then the span of S is all vectors that can be expressed $a_1v_1 + a_2v_2 + a_3v_3$ for some numbers a_1, a_2, a_3.

(This definition of span is imperfect. So that other things work, define $\mathrm{span}(\{\}) = 0_v$

Theorem. *By the way that span is defined, the span of every vector is guaranteed to be a subspace, because the conditions of the last section are met.*

Proof. The first condition is that 0_v is in the span. This is true because $0_v = 0$ times any vector in the span.
The next condition is that the sum of two elements in the span is also in the span. Take $S = \{v_1, v_2, \dots\}$ and two vectors u and w in the span. Then for some numbers a and b, $u = a_1v_1 + a_2v_2 + \dots$ and $w = b_1v_1 + b_2v_2 + \dots$. Then $u + w = (a_1 + b_1)v_1 + (a_2 + b_2)v_2 + \dots$, which is the sum of multiples of spanning vectors, so the sum is in the span.
The last condition is that multiples of vectors in the span are also in the span. The proof is very similar to the one above. $\qquad \square$

4.2 Examples

1. Take the vectors space of lists of numbers of size three, and take the set with two element $S = \{(1, 0, -1), (0, 1, -1)\}$ Elements in span(S) have the form $a(1, 0, -1) + b(0, 1, -1) = (a, b, -a - b)$, where a and b are numbers. Notice that this is the subspace discussed above of lists of legth three whose elements add up to zero. To prove this, show that every combination has elements that add up to zero, and that every list whose elements add up to zero can be written as a linear combination.

2. In the vector space of polynomials whose largest term includes x^2, take span$\{1, x, x^2, 1 + x + x^2\}$. Every vector in this vector space is in the span. However, there at least two ways to write every vector in the space. The vector $2 + 4x - 3x^2$ can be written as $2(1) + 4(x) - 3(x^2)$, or as $-3(1 + x + x^2) + 5(1) + 7(x)$. In fact, there are an infinite number of ways that each vector can be written.

4.3 Linear Independence

In the last example, there are more than one ways to write each element of the span because the last element of the spanning set is the sum of the first three elements or, stated another way, that the last element is in the span of the first three: $1(1) + 1(x) + 1(x^2) = (1 + x + x^2)$. A set of vectors which contains an element that can be written as the span of other vectors in the span is called **linearly dependent**, and the opposite, a set with elements that can't be written as the sum of multiples of other elements, is called **linearly independent**.

Here are two important facts about linear independence:

Theorem. *A set is guaranteed to contain a linearly independent subset.*

Proof. Take a linearly dependent set. Remove an element which can be written as linear combinations of other set elements. If the set is still linearly dependent, repeat. A set with only one element is linearly independent, so the process will eventually give a linearly independent set. Also, because the elements that were removed were linear combinations of the elements that are left, the span of the set and the span of the linearly independent part is the same. $\square$

Theorem. *Elements of the span of a linearly independent set can only be written as multiples of set elements in one unique way.*

Proof. Assume that this isn't true[3]. Take the linearly independent set $v_1, v_2, \ldots$. Say that there are two ways to write the same vector w: $w = a_1 v_1 + a_2 v_2 + \ldots$, and also $w = b_1 v_1, b_2 v_2 + \ldots$, where a_i and b_i are numbers, and not all $a_i = b_i$. Then $w - w = 0 = a_1 v_1 + a_2 v_2 + \cdots - (b_1 v_1 + b_2 v_2 + \ldots) = (a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \ldots$ Isolating v_1 gives $v_1 = -\frac{1}{a_1 - b_1}(a_2 - b_2)v_2 + (a_3 - b_3)v_3 + \ldots$. However, the set of v_i was said to be linearly independent, meaning that it

---

[3]'proof by contradiction'

shouldn't be possible to write v_1 as the sum of multiples of the other vectors. This contradiction means that there can't be two ways to write w. $\qquad\square$

5   Basis and Dimension

5.1   Basis

A **basis** of a space V is a set β where $\text{span}(\beta) = V$, and β is linearly independent. Each set has more than one possible basis, but each space has a basis, and the basis of each set is the same size.

Theorem. *Each space has a basis.*

Proof. Take the whole space. The span of the space is the space itself, by the first rules for vector addition and multiplication. The set might be linearly dependent, so apply the process from the first theorem in the section above, which will give a linearly independent subset which still generates the whole space. $\qquad\square$

The proof of the second fact is hard, technical, and not very interesting, but the result is interesting and important. It allows us to compare the sizes of vector spaces.

5.2   Dimension

The number of basis elements of a vector space is the **dimension** of the space. Because all bases of a space have the same number of elements, spaces only have one dimension. Dimension fits our intuitive understanding of the 'size' of a vector space: Subspaces have a smaller dimension than the spaces that contain them, and if two spaces have the same dimension, then there are one-to-one correspondences between the two spaces.

Theorem. *Subspaces have a dimension less than or equal to the dimension than the space that contains it.*

Proof. Take a basis of the subspace. The basis cannot contain more elements than the dimension of the containing space, because the basis elements of are elements of the containing space, so if the basis has more elements, it can't be linearly independent, because the basis of the containing space spans it. $\qquad\square$

The last fact about one-to-one correspondences will be discussed in a later section about transformations.

6   Transformations

Functions between vector spaces are exactly the same as functions between standard numbers. A function between vector spaces is a just a map which

takes some vectors to some others. Functions can output vectors in a different space than the input: The 'input' space is called the **codomain** and the 'output' space is the **domain**.

The functions most of interest to linear algebra are the **linear transformations**, because linear transformations preserve the structure of vector spaces. Any function T is a linear transformation if has the following properties:

1. For vectors v and w, $T(v + w) = T(v) + T(w)$.

2. For vector v and number a, $T(av) = aT(v)$.

These two simple requirements are enough to allow many strong, surprising conclusions about linear functions.

6.1 Examples

As an exercise, check that all of these really are linear.

1. In the two-dimensional plane, a reflection across the line $y = x$

2. A function from lists of size three to lists of size two that takes $(a, b, c) \mapsto (a, -a)$

3. A function from lists of size two to lists of size four that takes $(a, b) \mapsto (a - b, 0, 0, 0)$

4. A function from lists of size three to polynomials of degree four that takes $(a, b, c) \mapsto a + (a + b)x + (a + b + c)x^2 + bx^3 + (b + c)x^4$

5. A function from numbers in three-dimensional space that takes $a \mapsto$ the point with coordinates $(a, 2a, 4a)$

6.2 Appendix? Addendum? Anyway matrices

It's possible to arrange the representation of vectors and transformations to reduce all transformations to a simple, rote calculation. This section assumes that you know about matrices, and how they multiply.

The first step is to represent each vector with a one-column matrix. To do this, choose a basis, give the elements an order, and write each matrix as a combination of multiples of basis elements in the given order. In the vector space of lists of size three, choose basis $\beta = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. Then $(3, 1, -2.4) = 3(1, 0, 0) + 1(0, 1, 0) - 2.4(0, 0, 1)$. Next, fill in the rows of a matrix with the multiples from top to bottom, again in order.

So $[(3, 1, -2.4)]_\beta \mapsto \begin{pmatrix} 3 \\ 1 \\ -2.4 \end{pmatrix}$

Matrix addition and scalar multiplication act in the same ways vector addition and multiplication by numbers do. That means that this representation of a vector captures everything about it relevant to linear algebra.

The next step is to represent transformations. Each transformation will be a $n \times m$ matrix, where n is the dimension of the domain and m the dimension of the codomain. Each of the n columns is the column representation of what transformation to each element of the basis, in the same order.

As an example, take the space of lists of size three, and consider the linear function which reverses the order of the list. Taking the same basis as above, the function takes:

$$(1,0,0) \mapsto (0,0,1) \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$(0,1,0) \mapsto (0,1,0) \mapsto \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$(0,0,1) \mapsto (1,0,0) \mapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

So the matrix representation of this function is $[T]_\beta = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

Now, application of the function to a vector v is the same as the multiplication of the column representation by this matrix representation. $[v]_\beta * [T]_\beta$ gives the column vector $[T(v)]_\beta$. The action of linear transformations has been reduced to matrix multiplication. [4]

6.3   Range and Nullity

Given a transformation T, define two new sets as follows:

- The **range** is all vectors x such that there's some y that $x = T(y)$. The range is a subset of the codomain. In example 2 above, only lists whose elements add up to zero are in the range.

- The **null space** are all vectors x such that $T(x) = 0_v$. The null space is a subset of the domain. In example 3, list whose elements are all the same are in the null space.

Theorem. *The range of a linear transformation is a subspace.*

Proof. Need to check the three conditions: First, that 0_v is in the range. Taking the second condition of linear transformations with $a = 0 \implies T(0v) = 0T(v) \implies T(0_v) = 0_v$. 0_v is the image of 0_v, so it is in the range. Next, show that the sum of two things in the range is in the range. Take x and y in the range. There must be z and α so that $x = T(z), y = T(\alpha)$. So

---

[4]Some interesting implications of this: All matrices correspond to a linear transformation. Matrices are just a special case of lists of numbers, so they're a vector space. Therefore, there's a way to cosider linear functions as a vector space, with their own basis, dimension, and functions. As an excercise, try to figure out the meaning of addition and scalar multiplication on the vector space of linear functions.

$x + y = T(z) + T(\alpha) + T(z + \alpha)$. The proof that multiples of elements of the range are in the range is similar. $\square$

Theorem. *The null space of a linear transformation is a subspace.*

Proof. Already know from last proof that $T(0_v) = 0_v \implies 0_v$ is in the null space. Next, check that for x, y in null space $x + y$ in null space. $T(x + y) = T(x) + T(y) = 0_v + 0_v = 0_v \implies x + y$ in null space. Proof of last part similar. $\square$

6.4   The Rank-Nullity Theorem

Because the range and null space are vector spaces, they have their own bases and dimension. The rank-nullity theorem is an important result which relates these dimensions to the dimension of the domain.

For convience, let the **rank** of a transformation be the dimension of the range and **nullity** the dimension of the null space.

Theorem. *For any linear transformation, the rank plus the nullity is equal to the dimension of the domain.*

Proof. Take a linear transformation W with domain V and codomain W, with $\dim(V) = n$. Take a basis of the nullity $\{w_1, \ldots, w_i\}$. Notice that the nullity is i. Add enough vectors $\{v_1, \ldots, v_j\}$ so that $\{w_1, \ldots, w_i, v_1, \ldots, v_j\}$ is a basis for V, implying that $j + i = n$.

Next, show that $\{T(v_1), \ldots, T(v_j)\}$ is a basis for the range with j elements.

First, show that this set is linearly independent. By contradiction, suppose $T(v_1) = a_2 T(v_2) + \cdots + a_j T(v_j)$. By the rules of linear transformations, $T(v_1) - a_2 T(v_2) - \cdots - a_j T(v_j) = 0 \implies T(v_1 - a_2 v_2 - \cdots - a_j v_j) = 0 \implies v_1 - a_2 v_2 - \cdots - a_j v_j$ is in the null space. However, this is a contradiction, because these vectors are linearly independent with the basis of the null space. This contradiction means that $\{T(v_1), \ldots, T(v_j)\}$ is linearly independent.

Second, show that $\{T(v_1), \ldots, T(v_j)\}$ is a span of the range. For every element of the range $x = T(y), y$ is an element of V, so it can be written in terms of the basis elements: $x = T(a_1 w_1 + \cdots + a_i w_i + a_{i+1} v_1 + a_n v_j) = a_1 T(w_1) + \cdots + a_n T(v_j) = 0_v + a_{i+1} T(v_1) + \cdots + a_n T(v_j)$ because all w are in the nullity. This shows that all x in the range can be written as a combination of multiples of $\{T(v_1), \ldots, T(v_j)\}$. Therefore, that set spans the range.

The set spans the range, and is linearly independent. Therefore, the basis of the range has j elements, so the rank of T is j. From above, the nullity was i, and $j + i = n = \dim(V)$ $\square$

6.5   Injection, surjection, and bijection

To classify functions, consider their domain and codomain. An **injective** function takes every element of the domain to a different element of the codomain. An example is the linear function $(a, b, c) \mapsto (0, a, -b, 0, 3c)$. A **surjective** function has the same range and codomain – Every element of the codomain is the

image of at least one element of the range. The function $(a, b, c) \mapsto (a - c, b)$ is surjective. A **bijective** or one-to-one function is both injective and bijective. The function $(a, b, c, d) \mapsto \begin{pmatrix} a + d & a - d \\ b - c & b + c \end{pmatrix}$

There are a few facts that follow from the definitions here and the rank-nullity theorem worth knowing.

Theorem. *Take linear function T. If $N(T) = \{0_v\}$, T is injective.*

Proof. Take v_1 and v_2 so that $T(v_1) = T(v_2) = w$. $\implies T(v_1) - T(v_2) = T(v_1 - v_2) = 0$. Because only 0_v is in the nullity, $v_1 - v_2 = 0 \implies v_1 = v_2$. For each vector w in the codomain, there is only one unique v_1 with image w. $\square$

Theorem. *Take T linear function. If T is injective, $N(T) = \{0_v\}$*

Proof. Already know that 0_v in $N(T)$. Need to show that only 0_v in $N(T)$. Take a in $N(T)$. $T(a) = 0$, so for any v, $T(v) + 0_v = T(v) + T(a) = T(v + a) = T(v)$. T is injective, so $v = v + a \implies a = 0$ $\square$

Theorem. *Take linear function $T : V \to W$ (domain V and codomain W). If T is injective, $dim(V) \leq dim(W)$.*

Proof. Nullity$(T) = 0$, so rank$(T) = \dim(V)$ by rank-nullity theorem. rank(T) $\leq \dim(W)$, so $\dim(V) \leq \dim(W)$. $\square$

Theorem. *Take T surjective linear function from V to W. $dim(V) \geq dim(W)$.*

Proof. Range$(T) = W \implies$ rank$(T) = \dim(W)$. By rank-nullity theorem, $\dim(V) = $ nullity$(T) + \dim(W)$. nullity$(T) \geq 0$, so $\dim(V) \geq \dim(W)$. $\square$

Matrices, rank, and nullity

Recall that columns of a matrix representing a transformation represent the image of each element of the chosen basis. Therefore, the span of the columns is range of the transformation.

Proof. Take matrix representing a transformation T under basis $\beta = \{v_1 \dots v_n\}$ with columns $w_1 \dots w_m$. First, show every element x of $R(T)$ is in span$\{w_1 \dots w_n\}$. Every $x = T(a_1 v_1 + \dots + a_n v_n) = a_1 T(v_1) + \dots + a_n T(v_n) = a_1 w_1 + \dots + a_n w_n$. Next, show every y in the span is in $R(T)$: $y = a_1 w_1 + \dots + a_n w_n = a_1 T(v_1) + \dots + a_n T(v_n) = T(a_1 v_1 + \dots + a_n v_n)$. $\square$

So the rank of a transformation is number of linearly independent columns in any matrix representation. By the rank-nullity theorem, the nullity is the total number of columns minus the number of linearly independent ones, because the total number of columns is the dimension of the domain.

A matrix with a nullity of n has n column vectors which can be written as linear combinations of the other vectors. Each equation desciribing how one of the n vectors is the sum of the others corresponds to one of the elements of a basis for the null space of T.