

Linear Algebra

Karan Elangovan

March 27, 2021

Contents

1	Vector Spaces	1
1.1	Vector Space Axioms	1
1.2	Linear Independence, Spanning Sets and Bases	2
2	Linear Maps	5
2.1	Definitions	5
2.2	Quotient Spaces	6
2.3	Matrices	8
2.4	Operators	8
3	Eigenvalues and Eigenvectors	9
3.1	Invariant Subspaces	9
3.2	Upper Triangular Matrices	12

1 Vector Spaces

1.1 Vector Space Axioms

Definition 1 (Vector Space). A vector space over a field k is an ordered triple $(V, +, f)$, with $+$: $V^2 \rightarrow V$ and f : $k \times V \rightarrow V$ satisfying

1. $(V, +)$ is an abelian group
2. $1_k v = v$ for all $v \in V$
3. $a(bv) = (ab)v$ for all $a, b \in k$ and $vin \in V$
4. $(a + b)v = av + bv$ for all $a, b \in k$ and $vin \in V$
5. $a(v_1 + v_2) = av_1 + v_2$ for all $a \in k$ and $v_1, v_2 \in V$

Where we use adjacency to denote f .

We call the elements of V vectors, the elements of k scalars and f scalar multiplication.

Definition 2 (Subspace). A subspace of a vector space V over k is a subset $U \subset V$ which is also a vector space over k under the induced operations of addition and scalar multiplication.

1.2 Linear Independence, Spanning Sets and Bases

Definition 3 (Linear Combination). Let $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ be a finite¹ set of vectors in V . Then we define a linear combination of \mathcal{V} as a function $f : \mathcal{V} \rightarrow k$. To each linear combination corresponds a vector $f(v_1)v_1 + f(v_2)v_2 + \dots + f(v_n)v_n$ as a linear combination. Whether we are speaking of the actual vector corresponding to a linear combination or the choice of scalars is usually easily inferred from context.

Definition 4 (Span). We define the span of a finite set of vectors \mathcal{V} as the set of all of its linear combinations. If the span is V then we say \mathcal{V} is a spanning set.

¹ The definition may be extended for sets of infinite cardinality as well, with the restriction that all but finitely many of the choices made by f are 0, allowing us to sidestep issues of infinite sums. However most useful applications and results are with regards to the finite case, so we restrict our definitions and study to this case.

Theorem 1. Let \mathcal{V} be a finite set of vectors. Then the following are all equivalent

1. The only linear combination of \mathcal{V} equal to 0 has all of its scalars equal to 0.
2. For every v in the span of \mathcal{V} , the choice of scalars is unique.
3. For every $v \in \mathcal{V}$, v is not in the span of the remaining vectors in \mathcal{V} .

Proof. (2) is a stronger form of (1), hence (2) implies (1). Now assume (1). Say two linear combinations $a_1v_1 + a_2v_2 + \dots + a_nv_n$ and $b_1v_1 + b_2v_2 + \dots + b_nv_n$ are both equal to v , then we have $(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \dots + (a_n - b_n)v_n = 0$, which by (1) yields $(a_i - b_i) = 0$, giving $a_i = b_i$, hence the choice of scalars is unique, and so (1) implies (2).

Now assume (1). Say (3) is not true, then we may assume WLOG by re-indexing that $v_n = a_1v_1 + a_2v_2 + \dots + a_{n-1}v_{n-1}$, yielding a non-trivial linear combination that is 0, contradicting (1), hence (1) implies (3). Now assume (3), and say that (1) is not true. Then we have $a_1v_1 + a_2v_2 + \dots + a_nv_n = 0$ with some $a_i \neq 0$. So we may rearrange to yield $a_i = (-\frac{a_1}{a_i})v_1 + \dots + (-\frac{a_{i-1}}{a_i})v_{i-1} + (-\frac{a_{i+1}}{a_i})v_{i+1} + \dots + (-\frac{a_n}{a_i})v_n$, contradicting (3). Hence (3) implies (1). \square

Definition 5 (Linear Independence). We say a finite set of vectors \mathcal{V} is linearly independent iff it satisfies any of the equivalent conditions in theorem 1.

Theorem 2. The cardinality of any linearly independent set cannot exceed the cardinality of any spanning set.

Proof. Let $\mathcal{W} = \{w_1, w_2, \dots, w_m\}$ be an arbitrary linearly independent set and $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ an arbitrary spanning set.

As \mathcal{V} is spanning, w_1 belongs to its span. We claim that we may replace m of the v_i with distinct elements of \mathcal{W} to yield a spanning set. We

proceed by induction. Assume that the set

$$S = \{w_1, w_2, \dots, w_j, u_1, u_2, \dots, u_{n-j}\}.$$

where $j < m$ and u_i are distinct elements of \mathcal{V} is spanning. Then w_{j+1} must belong to the span of S . If $w_{j+1} \in S$, then it must equal one of the u_i , and so we may replace that u_i with w_{j+1} . Otherwise consider

$$S' = \{w_{j+1}, w_1, w_2, \dots, w_j, u_1, u_2, \dots, u_{n-j}\}.$$

which must be linearly dependent as w_{j+1} belongs to the span of the remaining vectors. We know that $\{w_{j+1}, w_1, w_2, \dots, w_j\}$ is linearly independent, and so if we keep adding u_i we eventually reach the first u_i in the span of

$$\{w_{j+1}, w_1, w_2, \dots, w_j, u_1, u_2, \dots, u_{i-1}\}.$$

which we know exists as otherwise S' would be linearly independent. Hence we may discard this u_i to leave the desired spanning set. The base case follows from \mathcal{V} being spanning. Hence we have an injection from $\mathcal{W} \rightarrow \mathcal{V}$ by which $v \in \mathcal{V}$ is replaced by $w \in \mathcal{W}$ in the construction. Hence $|\mathcal{W}| \leq |\mathcal{V}|$. \square

As a trivial consequence of this, we have that any two bases have the same cardinality as their cardinalities bound each other above and below as a basis is simultaneously a spanning set and a linearly independent set. So we may call this invariant quantity the "dimension" of the vector space. However there is a caveat, namely that this assumes that there is a basis at all. When we extend our definitions to include the idea of a linear combination of an infinite set as mentioned in footnote 1 it may be shown that every vector space has a basis (this is a trivial consequence of Zorn's lemma when we view a basis as a

maximal linearly independent set)

Definition 6 (Finite-Dimensional Vector Spaces). The vector space V is finite dimensional iff it has a basis. If V is finite dimensional, then we call the common cardinality of all of its bases the dimension of V , $\dim V$.

2 Linear Maps

2.1 Definitions

The structure of a vector space is given by the operations of addition and scalar multiplication, so those maps between vector spaces which preserve this structure will be of interest. In the context of linear algebra, we call a homomorphism a linear map.

Definition 7 (Linear Map). Let V and W be k -vector spaces. The function $T : V \rightarrow W$ is a linear map iff it satisfies

1. $T(u + v) = T(u) + T(v)$ for all $u, v \in V$.
2. $T(\lambda v) = \lambda T(v)$ for all $\lambda \in k$ and $v \in V$.

We denote the set of all linear maps from V to W by $\mathcal{L}(V, W)$.

Theorem 3. Let $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ be a basis of V . Then for any choice of $w_1, w_2, \dots, w_n \in W$, there is a unique linear map $T \in \mathcal{L}(V, W)$ that maps each v_i to w_i .

Proof. Consider the mapping from $V \rightarrow W$ defined by

$$T(a_1 v_1 + a_2 v_2 + \dots + a_n v_n) = a_1 w_1 + a_2 w_2 + \dots + a_n w_n.$$

This is well-defined as \mathcal{V} is a basis and hence each vector has a unique representation as a linear combination. Homogeneity and additivity are trivially

verifiable. Hence there does exist such a map. The uniqueness follows from expressing each vector as a linear combination of \mathcal{V} , from which the requirements of a linear map and our hypothesis leave only one possible choice for the image.

□

A trivial but noteworthy corollary is that if V and W are finite-dimensional and of equal dimension they must have a bijective linear map between them, meaning they are isomorphic. This means that for finite dimensional vector spaces, the dimension totally determines the structure. So we may treat k^n as the canonical vector space of dimension n .

Definition 8 (Kernel). Let $T \in \mathcal{L}(V, W)$ be a linear map. Then the kernel of T , $\ker T$ is the pre-image of 0.

Similarly to group homomorphisms, the kernel of a linear map is a subspace, and the triviality of the kernel is equivalent to injectivity, and there is also as one would expect an analogue to the isomorphism theorem for groups which we will be able to state and prove once having defined the notion of the quotient of a vector space and its subspace.

2.2 Quotient Spaces

Consider a vector space V and a subspace U . We wish to define a vector space V/U that is conceptually the vector space V , but in which we consider vectors which differ only by some vector in U as identical. More precisely V/U is the set of equivalence classes under the equivalence relation $v \sim u \iff (v - u) \in U$, with addition and scalar multiplication defined on coset representatives. These operations may be shown to be well-defined and to yield a vector space, we omit the proofs as they are essentially identical in logic to showing the cosets of a normal subgroup form a group.

Definition 9 (Quotient Space). The quotient space V/U is the set of cosets of U in V , with the operations of addition and scalar multiplication defined on coset representatives.

Theorem 4 (Fundamental Theorem of Linear Maps). Let $T \in \mathcal{L}(V, W)$ be a surjective linear map. Then $V/\ker T \cong W$.

Proof. We define the map $T' : V/\ker T \rightarrow W$ by

$$T'(\ker T + v) = T(v).$$

T' is well-defined as no matter which coset representative we choose, their difference will be in the kernel of T , and so will have no effect on the image by the additivity of T . T' is also clearly linear as T is. If $\ker T + v \in \ker(T')$, then $T(v) = 0$, and so $\ker v + v = \ker v + 0$, hence the kernel of T' is trivial and so T' is injective. The surjectivity of T' follows from the surjectivity of T . Hence $V/\ker T$ and W are isomorphic by T' . \square

Theorem 5. Let V be finite-dimensional, and U a subspace of V . Then $\dim(V/U) = \dim V - \dim U$.

Proof. Let u_1, u_2, \dots, u_n be a basis of U . Let us extend it into a basis for V to $u_1, u_2, \dots, u_n, v_1, v_2, \dots, v_m$. Hence $\dim(V) - \dim(U) = (n + m) - (n) = m$.

We claim that $\mathcal{V} = \{U + v_1, U + v_2, \dots, U + v_m\}$ is a basis of V/U .

$$\begin{aligned} & \{U + u_1, U + u_2, \dots, U + u_n, U + v_1, U + v_2, \dots, U + v_m\} \\ &= \{U + 0, U + v_1, U + v_2, \dots, U + v_m\} \end{aligned}$$

spans V/U , and hence so does \mathcal{V} , as the 0 vector doesn't affect the span. Say we have a 0 linear combination of \mathcal{V} , $a_1(U + v_1) + a_2(U + v_2) + \dots + a_m(U + v_m) =$

0, so we have $a_1v_1 + a_2v_2 + \dots + a_mv_m \in U$, so we have $b_1, b_2, \dots, b_n \in k$ such that $a_1v_1 + \dots + a_mv_m = b_1u_1 + \dots + b_mu_m$ as $\{u_1, \dots, u_m\}$ spans U . Hence by the linear independence of the entire basis of V , we have $a_1 = \dots = a_m = b_1 = \dots = b_n = 0$. Hence there are no non-trivial 0 linear combinations of \mathcal{V} . Hence \mathcal{V} is a basis as we claimed, and so $\dim(V/U) = m = \dim(V) - \dim(U)$. \square

2.3 Matrices

As we have established, a linear map $T \in \mathcal{L}(V, W)$ is uniquely identified by where it maps any given basis \mathcal{V} of V , and to every possible choice of images of \mathcal{V} there corresponds a linear map. This motivates us to define a matrix as a way to represent linear maps.

Definition 10 (Matrix). Let $T \in \mathcal{L}(V, W)$, and $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ and $\mathcal{W} = \{w_1, w_2, \dots, w_m\}$ be bases of V and W , respectively. Then we define the matrix of T with respect to \mathcal{V} and \mathcal{W} , $\mathcal{M}(T, \mathcal{V}, \mathcal{W})$ as an array with m rows and n columns, such that column j gives the scalars of the linear combination of \mathcal{W} which corresponds to $T(v_j)$.

We define addition, multiplication (when the corresponding linear maps are composable) and scalar multiplication on matrices such that the resulting matrix corresponds to the resulting linear map when the corresponding linear maps of the original matrices are combined. Symbolically, $\mathcal{M}(S) + \mathcal{M}(T) = \mathcal{M}(T + S)$, $\lambda\mathcal{M}(T) = \mathcal{M}(\lambda T)$, and $\mathcal{M}(S)\mathcal{M}(T) = \mathcal{M}(S \circ T)$.

2.4 Operators

We consider the special case of linear maps whose domain and range are identical, that is, the endomorphisms of a vector space. The primary characteristic that leads to their richer theory is that any operator may be composed with itself, so we may consider powers of an operator, and hence polynomials of an

operator.

Definition 11 (Operator). An operator is a linear map whose domain is the same as its range. We denote the operators of V as $\mathcal{L}(V)$.

Theorem 6. For an operator T on a finite-dimensional space V , injectivity is equivalent to surjectivity.

Proof. Let V have dimension n . We have by the fundamental theorem of linear maps, $V/\ker T \cong T$.

First assume that T is injective, then the kernel is trivial and hence we have $V \cong T$, and so $\dim T = n$, and as T is a subspace of V , we have that $T = V$, meaning T is surjective.

Now assume that T is surjective, then we have $V/\ker T \cong V$, and so $\dim \ker T = 0$, hence the kernel is trivial and so T is injective. \square

3 Eigenvalues and Eigenvectors

3.1 Invariant Subspaces

We may better understand an operator on V by instead considering how it acts on each subspace in some direct sum decomposition of $V = U_1 \oplus \dots \oplus U_n$. However T restricted to U_i , $T|_{U_i}$ may not be an operator. It is only an operator when it maps the subspace into itself. This motivates our definition of invariant subspaces.

Definition 12 (Invariant Subspace). We call the subspace U invariant under $T \in \mathcal{L}(V)$ iff $T(U) \subset U$, or equivalently $T|_U$ is an operator.

The simplest case is when the subspace has dimension 1, in which case it is the span of some non-zero vector v , and so it must be the case that $T(v) = \lambda v$

for some $\lambda \in k$. When this is the case we call v an eigenvector and λ its corresponding eigenvalue. Whilst every eigenvector may of course only correspond to a single eigenvalue, the converse does not hold, indeed the eigenvectors of a specific eigenvalue (which, alongside 0, trivially form a subspace) need not even have dimension 1.

Definition 13. Let $T \in \mathcal{V}$. Then $\lambda \in k$ is an eigenvalue iff any of the following equivalent criteria hold

1. We have some non-zero $v \in V$ such that $T(v) = \lambda v$
2. $T - \lambda I$ is not injective.

For any $\lambda \in k$, we define the eigenspace $E(\lambda, T)$ as the kernel of $T - \lambda I$.

Theorem 7. If v_1, v_2, \dots, v_n are eigenvectors with distinct corresponding eigenvalues, then $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ is linearly independent.

Proof. Assume the contrary, and so we have some j such that $\{v_1, v_2, \dots, v_j\}$ is linearly independent and $\{v_1, v_2, \dots, v_{j+1}\}$ is linearly dependent. Hence we have for some choice of $a_1, a_2, \dots, a_j \in k$ that $v_{j+1} = a_1 v_1 + a_2 v_2 + \dots + a_j v_j$. We have by scalar multiplication by λ_{j+1} and by application of T that

$$\begin{aligned}\lambda_{j+1} v_{j+1} &= \lambda_{j+1} a_1 v_1 + \lambda_{j+1} a_2 v_2 + \dots + \lambda_{j+1} a_j v_j \\ \lambda_{j+1} v_{j+1} &= \lambda_1 a_1 v_1 + \lambda_2 a_2 v_2 + \dots + \lambda_j a_j v_j\end{aligned}$$

Hence we have

$$a_1(\lambda_{j+1} - \lambda_1)v_1 + a_2(\lambda_{j+1} - \lambda_2)v_2 + \dots + a_j(\lambda_{j+1} - \lambda_j)v_j = 0$$

The eigenvalues are all pairwise distinct, and hence by the linear independence of $\{v_1, v_2, \dots, v_j\}$ we have that $a_1 = a_2 = \dots = a_j = 0$. However this gives $v_{j+1} = 0$, which cannot be. Hence \mathcal{V} is linearly independent. \square

As these eigenvalues are helpful in understanding the behaviour of an operator, we should ask for which operators there even exists an eigenvalue. In the most general case, if the operator is non-injective, then 0 will be an eigenvalue. Otherwise we restrict our attention to finite dimensional \mathbb{R} and \mathbb{C} vector spaces. An arbitrary operator on a finite dimensional \mathbb{R} vector space need not have any eigenvalues (for example an order 4 turn about the origin), however as the next theorem shows, the answer is in the affirmative for finite dimensional \mathbb{C} vector spaces.

Theorem 8. Let T be an operator over a finite dimensional, non-trivial, \mathbb{C} vector space, V . Then T has an eigenvalue.

Proof. Let $\dim V = n$. Choose an arbitrary non-zero $v \in V$. Then $\mathcal{V} = \{v, T(v), \dots, T^n(v)\}$ is linearly dependent, so we have some non-trivial zero linear combination of \mathcal{V} . Let the scalars be a_0, a_1, \dots, a_n and j be the last non-zero index. j must exist as the combination is non-trivial, and $j \geq 1$, else we'd have that $v = 0$. So we have

$$(a_0I + a_1T + a_2T^2 + \dots + a_jT^j)v = 0$$

$j \geq 1$, hence by the Fundamental Theorem of Algebra, we may factorise the polynomial into a product of linear factors.

$$(T - \lambda_1I)(T - \lambda_2I) \dots (T - \lambda_jI)v = 0$$

Hence not all of the linear factors may be bijective, and so at least one, $(T - \lambda_jI)$, is non-injective, injectivity is equivalent to bijectivity over finite dimensional vector spaces. This means that λ_j is the desired eigenvalue. \square

3.2 Upper Triangular Matrices

Definition 14 (Upper Triangular Matrix). We say a square matrix is upper triangular iff all entries under the diagonal are 0.

If we consider an operator $T \in \mathcal{L}(V)$, we may express the condition of upper-triangularity of $\mathcal{M}(T, \mathcal{V})$ in terms of the operator. Say $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$, then it is trivial to see that upper triangularity is equivalent to each $T(v_j)$ belonging to the span of $\{v_1, v_2, \dots, v_j\}$.

Theorem 9. Let T be an operator over a finite-dimensional \mathbb{C} vector space. Then there exists a basis \mathcal{V} such that $\mathcal{M}(T, \mathcal{V})$ is upper-triangular.

Proof. Assume by induction that we have a linearly independent set $\mathcal{V} = \{v_1, v_2, \dots, v_j\}$ such that each $T(v_i) \in \text{span}\{v_1, v_2, \dots, v_i\}$, with $j < n$.

Let $U = \mathcal{V}$. So V/U is a non-trivial \mathbb{C} vector space, and U is an invariant subspace under T . So we have an eigenvector $U + v_{j+1}$ of the operator T/U , so $(T/U)(U + v_{j+1}) = U + \lambda v_{j+1}$.

So if we let v_{j+1} be a coset representative of $U + v_{j+1}$ then we have $T(v_{j+1}) = a_1 v_1 + a_2 v_2 + \dots + a_j v_j + \lambda v_{j+1} \in \text{span}\{v_1, v_2, \dots, v_{j+1}\}$ and so we have $\mathcal{V} \cup \{v_{j+1}\}$ as a set of $j+1$ vectors that satisfies the property. So by induction we have the desired basis. \square