Linear Algebra

Purnima Tiwari

MLS01
Madhava Lecture Series, Theoretical Nexus

Contents

1. A note.

These notes have been drafted alongside the content of lectures. However, you might find a few additional things. The definition of a field and a vector space have been presumed to be of knowledge to the reader. We begin with the basic concepts and build upon them.
Happy Reading!

2. Part 1

Linear Algebra deals mainly with Vector Spaces, and Linear Transformations or Operators, where *Linear* refers to the preservation of vector addition and scalar multiplication under transformation.

Fields and Vector Spaces.

The definition of a *field* deals with a set of axioms that a non-empty set of elements must follow to be classified as a field. And the similar applies for the definition of a vector space, that is, if a set follows the notions of a usual vector or a linear space, then it is classified as a vector space. I will take the liberty to skip mentioning the axioms.

Now, the term *space* with respect to the linearity property, refers to the *freedom of movement.* When talking about vector spaces, the freedom of movement applies since, scaling, addition and multiplication etc., lead to the *spanning* of the vector space itself.

An important observation would be to notice that the vector space V is a composite object, that is defined over the field of scalars $\mathbb{F}$, consisting of two operations, which must satisfy a set of properties. Hence a vector space is usually defined over a field, denoted as $V(\mathbb{F})$.

Definition 1. *A* subfield *is defined as a subset of the field that also forms a field under the same set of axioms and the usual operations.*

An easy exercise left up to the reader would be to examine how the set of all complex numbers $(\mathbb{C})$ form a field and the set of all rational numbers $(\mathcal{Q})$ form a subfield of $(\mathbb{C})$. However, any other set of numbers which do not contain every rational number, will not form a subfield of $(\mathbb{C})$. Hence, the set of all positive integers $(\mathbb{Z}^+)$, the set of all integers $(\mathbb{Z})$, and the set of all natural numbers (N) do not form a subfield of $(\mathbb{C})$.

Definition 2. *A* subspace *is defined as the subset of the vector space over a field, say $\mathbb{F}$, along with the operations of vector addition and scalar multiplication. A good exercise would be to observe that a* *non-empty subset W of a vector space V, is a subspace, iff for each pair of vectors α and β in W, and each scalar c in $\mathbb{F}$, the vector $c\alpha + \beta$ is again in W.*

Note. *To imagine the concept of a vector space, one can think of a differential equation, such that the sum of two of its solutions results in the third solution, and multiplying a solution with a scalar, results in another solution. Now, these solutions can be thought of as vectors and the set of solutions as vector spaces.*

A good exercise would be to observe that, for a vector space $V(\mathbb{F})$, the intersection of any collection of subspaces of V will form a subspace of V.

Note. *To get things more intuitively, if one is familiar with the concept of group, then one can define* commutative group *as a set of vectors in a vector space, together with the operation of vector addition. Conversely, we could take an Abelian group and define a scalar multiplication on it to turn it into a vector space.*

Definition 3. *Consider a set of vectors V, such that we can create several linear combinations using all the vectors in V. Now, these linear combinations would 'make up' a linear space. This 'making up' of a linear space, is what is called as* spanning *the linear space, or the* span of the linear space.
Hence, it becomes easy to understand that a Linear Span of a set of vectors V, forms a Linear Space.

Consider a set of n vectors $x_1, x_2, \cdots, x_n$, and a linear space S, such that $x_1, x_2, \cdots, x_n \in S$. Then linear span of S will be expressed as, span$(x_1, ..., x_n)$= set of all linear combinations.

A good exercise would be to observe that a linear span forms the smallest subspace that contains a set of vectors.

Definition 4. *Let V be a vector space over $\mathbb{F}$, then a subset S of V is said to be* linearly dependent *if there exist distinct vectors $x_1, x_2, \cdots, x_n \in S$, and scalars $a_1, a_2, \cdots, a_n \in \mathbb{F}$, not all of which are zero, such that, $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$.*
A set which is not linearly dependent is said to be linearly independent.

I do agree that this might sound confusing at first, so to elaborate, when the set of vectors can be represented in the form of a linear combination, equal to zero, such that the vectors and the scalars are not all zero, then a relation of dependency establishes between the vectors and the scalars, and hence the vectors are referred to as linearly *dependent.* Now, if the set can be represented in the same linear combination, but now the scalar coefficients are all zero, then that would imply that the vectors themselves will not be dependent, and this is the set of linearly *independent* vectors.

A few simple observations, based on the definition would be:

(1) Any set which contains a linearly dependent set is linearly dependent.
(2) Any set which contains the zero vector, is always linearly dependent. $[\because a_i \cdot 0 = 0]$
(3) Any subset of a linearly independent set is linearly independent.
(4) A set of vectors is linearly independent, iff each finite subset of the set is linearly independent.

Definition 5. *For a vector space V, a basis is defined as a collection of a set of linearly independent vectors in V, which span the space V.*

Definition 6. *A standard basis for a field $\mathbb{F}^n$ can be expressed as, $S = (e_1, e_2, \cdots, e_n)$, where $e_1 = (1, 0, 0, \cdots, 0), e_2 = (0, 1, 0, \cdots, 0), \cdots, e_n = (0, 0, 0. \cdots, 1)$.*

Definition 7. *If V is a finite-dimensional vector space, then an* Ordered Basis *for V is a finite sequence of vectors which is linearly independent and spans V.*

Definition 8. Dimension *is a cardinal number; it is the cardinality of the basis. Hence, if a basis has n elements, then the vector space will be classified as* finite-dimensional, *with dimension n; and if the basis has an infinite number of elements, then the vector space is said to have* infinite-dimension.

Now, according to the above mentioned definition, if span$(v_1, v_2, \cdots, v_n) = V$, then V is finite-dimensional[1].

[1] n-dimensional, where n is finite.

An interesting exercise would be observe that the 'space' of polynomials is an infinite-dimensional vector space, since for a given polynomial of degree n, there exists a polynomial of degree $(n+1)$ in the space.

In a finite-dimensional vector space V, every non-empty linearly independent set of vectors is part of a basis.

Note. *Consider V to be a vector space over the filed $\mathbb{F}$, then the zero subspace of V, i.e., the set that contains only the zero vector will be spanned by the vector zero itself. But, $\{0\}$ is a linearly dependent subspace.*

Revisiting the definition of a linearly independent subspace, a set of vectors is classified as linearly independent iff the only possible linear combination resulting in zero, consists of all the scalars being equal to zero, and in the case of a zero subspace, the scalars 'need not be' zero, which establishes the zero subspace as a linearly dependent one.

Now, since the basis needs to have a non-zero element, but the zero subspace is spanned by the empty set[2] *$\{0\}$, hence a zero subspace is said to have zero dimension.*

Definition 9. *If $\mathbb{F}$ is a field of scalars, then it may be possible to add the unit 1 to itself a finite number of times and obtain zero, such that we can express it as,*

$$1 + 1 + 1 + ... + 1 = 0$$

If this happens in $\mathbb{F}$, then the least positive integer n for which the sum of n 1's is zero, is called the characteristic *of the field $\mathbb{F}$. And if this does not happen in $\mathbb{F}$, then it is referred to as a field of* characteristic zero.

Another definition of the characteristic of a field would be, 'the smallest natural number n, such that if we add the multiplicative identity n number of times, we obtain the additive identity'. Here, the 'multiplicative identity' and the 'additive identity' for scalars, can be considered as some *fancy terms* for 1 and 0, respectively.

The characteristic of a field becomes significant when dealing with algebraic structures (consider a group or a ring). For example, the characteristic of the field $\mathcal{Z}_2$ will be 2.

Note. *The characteristic of a field can either be* 0*, or a prime number, generally denoted as p.*

Definition 10. *When two vectors are perpendicular to each other in a vector space, they are referred to as* Orthogonal. *Mathematically, the dot product of two orthogonal vectors is zero.*

Orthogonality can be expressed as, $v_i \cdot v_j = 0$ for $i \neq j$.

Another way of expressing orthogonality would be say that, for two vectors v_i, v_j in the vector space V, then $v_i \perp v_j$ if $< v_i, v_j >= 0$, i.e., the inner product of the two orthogonal vectors is always zero.

Definition 11. *Consider a vector space V, and let S be a subset of V then,*
$S^{\perp} = \{$set of all elements $v \in V$ which are perpendicular to all the elements of $S\}$

This implies that, for all $s \in S, < v, s >= 0$.

$S^{\perp}$ is a subspace of V, and is referred to as the orthogonal space *of S.*

Consider the spaces $U = span\{(0,0,0)\}$, $V = span\{(1,0,0)(0,1,0)\}$, and $W = span\{(1,1,1)\}$.

Now, it must be seemingly observable that U is orthogonal to V as well as W. In fact, U will be orthogonal to every space.
Considering V and W, we realize that these two spaces will not be orthogonal to each other. But, if we introduce another space, say $S = span\{(0,0,1)\}$, then V and S, will be orthogonal to each other. Another good observation would be to see that the direct sum of V and S, will make up $\mathbb{R}^3$, represented as $V \bigoplus S = \mathbb{R}^3$.

Definition 12. *When two orthogonal vectors are also of unit length, they are referred to as* Orthonormal vectors.

Normality can be expressed as, $v_i \cdot v_i = 1$, for all i.

[2]We see that the set has no **non-zero** element.

Definition 13. *A linear map or a linear transformation is a functionsuch that, given two vector spaces V and W, such that, $f(v) = w$, where $v \in V$ and $w \in W$.*

Consider two vector spaces $V(\mathbb{F})$ and $W(\mathbb{F})$, with a linear mapping on them, defined as $T : V \to W$ $T(c\alpha + \beta) = c\ T(\alpha) + T(\beta)$ given any $\alpha, \beta \in V$ and for all $c \in \mathbb{F}$.

This function takes an element from the vector space V, and maps it to another element, lying in the vector space W.

•To eliminate any confusion, the terms 'linear map' and 'linear transformation' are synonymous to each other.

Now, consider a linear map $f : V \to W$ then,

(1) Image of the map forms a subspace of W is mathematically defined as,
$Img(f) = \{f(v) = w : w \in W\}$
Image consists of all the elememts in the vector space W, which have a pre-image in V.
(2) Kernel of the map is mathematically defined as,
$ker(f) = \{f(v) = 0 : v \in V\}$, and it forms a subspace of V.
Kernel consists of all the elements in the vector space V, which are mapped to the additive identity of the vector space W, which is generally represented by 0.

For those familiar with abstract algebra, the above mentioned terms of *image* and *kernel* will sound familiar and the concept of a linear map seems to resemble **homomorphism,** i.e., because linear transformations are Homomorphisms between vector spaces.

Every vector space is an Abelian group under addition, so the general theorems about homomorphisms and isomorphisms apply, and additional results, due to the second operation.

Then, $f : V \to W$ represents a linear map, and $dim\ V = dim\ kerf + dim\ Imgf$

Note. *A few points to keep in mind, regarding the duality bvetween linear maps and homomorphisms are,*

(1) *A linear map is injective, iff $ker = \{0\}$*
(2) *If $Img(f) = W$, then f is surjective, i.e., f is onto.*
(3) *f has an inverse iff it bijective, i.e., both injective and surjective.*
(4) *A linear map which has an inverse (or is invertible), is called an **Isomorphism**.*
(5) *Like every bijective function, f is non-invertible when there exists an element in the co-domain for which no pre-image exists.*

Some common examples of linear transformations are differnetiation, and integration.

A word on the singular and non-singular linear transformations

• Non-singular linear transformations have the $kerf = \{0\}$, i.e., only the additive identity of the first vector space is mapped to the additive identity of the another. These preserve linear independence. Consider T to be a linear transformation, such that for $T_\gamma = 0,\ \gamma = 0$

• Singular transformations are otherwise, hence the $kerf \neq \{0\}$, that is to say that the kernel will not be a singleton set zero.

• Hence for $T : V \to W$, T is non-singular iff T carries each linearly independent subset of V onto a linearly dependent subset of W.

Definition 14. *Consider V and W be two vector spaces over the field $\mathbb{F}$, and let T be a linear transformation from V into W, then the* null space *of T is the set os all vectors α in V such that $T\alpha = 0$.*

That is, for $T : V \to W$, null space of T $(N_T) = \{\alpha \in V : T\alpha = 0\}$

• Since we are familiar with the term *kernel,* we must be able to observe that,

Null space of a linear map= kernel of the map

Definition 15. *Consider a linear transformation from the vector space V, to another vector space W, $T : V \to W$, then the rank of the transformation T would be defined as the dimension of its image.*

Now, for the same transformaton T, the nullity of T is defined as the dimension of its kernel.

$$rank(T) = dim(Img(T))$$
$$nulltiy(T) = dim(ker(T)) = dim(N_T)$$

When the vector space V is finite dimensional, the rank-nullity theorem establishes that

$$rank(T) + nullity(T) = dimV$$

Note. *For a linear transformation T, if T^{-1} exists, then both T and T^{-1} are linear transformations.*

Definition 16. *A linear operator is a linear transformation from a vector space into itself.*

$$T : V \to V, \textit{is a linear operator.}$$

Definition 17. *Consider three vector spaces $V(\mathbb{F}), W(\mathbb{F})$ and $Z(\mathbb{F})$, and two linear transformations T and U, defined as*

$$T : V \to W$$
$$U : W \to Z$$

then, $TU : V \to Z$, this is the composition of two linear mappings.

Now, if the three vector spaces are equal, i.e. $V = W = Z$, then TU will also be defined in this case, where UT is a linear operator on V.

Definition 18. *A Linear Functional is a form of a linear transformation, such that, if V is a vector space over the field $\mathbb{F}$ (the field of scalars), then a linear map defined from V to F,*

$$f : V(\mathbb{F}) \to \mathbb{F}$$

is said to be an linear functional on V. The general form of the functional can be expressed as,

$$f(c\alpha + \beta) = c\ f(\alpha) + f(\beta)$$

for all $\alpha, \beta \in V$, and $c \in \mathbb{F}$.

A few points to note about the linear functionals are,

- A zero functional sends everything in the vector space V to zero (the scalar).
- The linearity principles apply.
- Linear functionals have a structure, similar to that of vector spaces.

Note. ***Trace*** *is a linear functional.*

Let $n \in \mathcal{Z}^+$ and consider $\mathbb{F}$ as the field of scalars. Consider a matrix A of order $n \times n$, with entries in $\mathbb{F}$, then the trace of A is a scalar.

Trace is mathematically defined as,

$$\text{Tr } A = A_{11} + A_{22} + \cdots + A_{nn}$$

Trace is a linear functional on the matrix space $\mathbb{F}^{n \times n}$ and can be expressed as,

$$\text{Tr } (cA + B) = \sum_{i=1}^{n} (cA_{ii} + B_{ii})$$
$$= c \sum_{i=1}^{n} A_{ii} + \sum_{i=1}^{n} B_{ii}$$
$$= c\ \text{Tr}\mathcal{A} + \text{Tr } B$$

Definition 19. *If V is a vector space, the collection of all the linear functionals forms a vector space in a natural manner.*

$L(V, \mathbb{F})$ denotes the collection of all the linear functionals on the vector space V over the filed $\mathbb{F}$. This collectionw hich now forms a space, is referred to as the dual space of V and is represented as V^*.

$$V^* = L(V, \mathbb{F})$$

Note. *If V is finite-dimensional, then this implies that so is V^* (the dual space of V), and*

$$dimV = dimV^*$$

.

- For a vector space V, let $\mathcal{B}$ denote the basis, such that,

$$\mathcal{B} = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$$

Now, there will precisely be a unique linear functional, say f_i for each i on V such that

$$f_i(a_j) = \delta_{ij}$$

From $\mathcal{B}$, we obtain a set of n independent linear functionals. Now, these linear functionals will also be linearly independent.

From the above, we conclude that $f_1, f_2, \cdots, f_n$ are n linearly independent functionals and since V^* (the dual space of the vector space of V) has the same dimension as V, i.e., n, then it must be that the basis of V^* will be

$$\mathcal{B}^* = \{f_1, f_2, \cdots, f_n\}$$

Hence, $\mathcal{B}^*$, which forms the basis of the dual space (V^*), is known as the **dual basis** of $\mathcal{B}$.

Definition 20. *In a vector space V of dimension n, a subspace of dimension $(n-1)$ is called a* hyperspace (V^ω).
These are defined as subspaces of codimension 1.

Note. (1) *Every hyperspace is the null space of a linear functional.*
(2) *A hyperspace is a maximal, proper subspace of V.*

A few points to note.

(1) For V being a vector space and V^ω being the hyperspace of V, by definition we have

$$dimV = n$$

$$dimV^\omega = n - 1$$

(2) Let f be a non-zero[3] linear functional, defined as

$$f : V \to \mathbb{F}$$

, then

$$rankf = 1$$

Since f forms a non-zero linear subspace of $\mathbb{F}$, the field of scalars, which itself is one-dimensional.

(3) Dimension of the null space of a linear functional will then be given as

$$dimN_f = dimV - 1$$

(using rank-nullity theorem)

Hence, if f is a non-zero linear functional, then the null space of f is a hyperspace in V, and every hyperspace in V is the null space of a (not unique) non-zero linear functional on V.

Definition 21. *Let V to be a vector space over the field $\mathbb{F}$, and let W be a subset of V, then V^* denotes the dual space of V, and annihilator of W in V^*, denoted as $W^\perp$ or W°, is the set of all linear functionals f on V, such that $f(w) = 0$, for every $w \in W$.*

A few points to note.

(1) W° is a subspace of V^*, even if W is not a subspace of V.
(2) If W is the set consisting of the zero vector alone, then $W^\circ = V^*$.
(3) If $W = V$, then W° is the zero subspace of V^*.
(4) There exists a natural isomorphism between W° and $(V/W)^*$.

[3] This implies that at least one image of an element of the vector space V will be non-zero.

Note. *Consider V to be a finite-dimensional vector space over the field $\mathbb{F}$, and let W be a subspace of V, then*

$$dim\ W + dim\ W^{\circ} = dim\ V$$

Definition 22. *Let V be a finite-dimensional vector space over the field $\mathbb{F}$, V^* be the dual space of V. We have already established that V^* forms a vector space, then we can also take the dual of the dual space, that is V^{**}, called the* ***double dual.***

For each vector α in V, we have $\alpha \mapsto f(\alpha)$ is a function, where α is an argument.

For each element $f(\alpha)$ in V^*, we have $\phi(f) = f(\alpha)$ is a functional, where α is a parameter.

- *If f is a linear function from V to the underlying scalar field, then the above linear maps are* dual *to each other.*

Note. *The vector space V and its double dual V^{**} have a natural isomorphism, whereas V and V^* do not have a natural isomorphism. In the latter case, the choice of basis determines or rather establishes the nature.*

Definition 23. *Let V be a vector space and W be a subspace of V such that a linear map T is defined fromV to W.*

$$T : V \to W$$

Now the dual map or transpose of T is defined as,

$$T^* : V^* \to W^*$$
$$T^*g = gT$$

for every $g \in W^*$, the dual map T^* sends a linear functional g on W^* to the composition gT, which is a linear functional on V.

Definition 24. *Let a vector space V with a linear transformation T_u from V onto itself defined as,*

$$T_u : V \to V$$
$$T_u(v) = v + u$$

In T_u, every v in the vector space V is 'translated' by a factor of u within the vector space V, then this is said to be a linear translation of V by u.

Definition 25. *Let a vector space V, and consider a linear transformation T defined from V onto itself, then a subspace $W \subseteq V$ is called an invariant subspace for T or T-*invariant, *if T transforms any vector $v \in W$ back into W.*

Hence, for every $v \in W$, $T(v) \in W$.

$$\Rightarrow TW \subseteq W$$

Note. *A T-invariant translation is an endomorphism*[4].

[4] A map from a vector space onto itself.

3. Part 2

To be updated soon.