Notes: Background on linear algebra

Rajat Mittal

IIT Kanpur

This is a small note on basic concepts in linear algebra which will be used in the course. These definitions are taken from Gilbert Strang's book "Linear algebra and its applications". For a detailed introduction to these concepts, please refer to Strang's book or any other elementary book on linear algebra.

We will be intereseted in vector space $\mathbb{R}^n$, but these concepts are valid for general vector spaces.

1 Vector space

A vector space is a set of elements closed under addition and scalar multiplication (all linear combinations). In other words, V is a vector space if for any $x, y \in V, \alpha \in \mathbb{R}$, $x + y, \alpha x \in V$ too.

There is a more formal definition with axioms about the binary operations and identity element. But the one above will provide the intuition for us. The most common examples for a vector space are $\mathbb{R}^n$, $\mathbb{C}^n$, space of all $m \times n$ matrices and space of all functions.

A subspace is a subset of a vector space which is also a vector space and hence closed under addition and scalar multiplication. A *span* of a set of vectors S is the set of all possible linear combinations of vectors in S. It forms a subspace.

Exercise 1. Give some examples of subspace of $\mathbb{R}^n$.

2 Linear independence

To understand the structure of a vector space, we need to understand how can all the elements be generated. From the definition, the concept of linear dependence/independence comes out.

Given a set of vectors $v_1, \cdots, v_n \in V$, they are *linearly dependent*, if vector 0 can be expressed as a linear combination of these vectors.

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0, \exists i, \alpha_i \neq 0.$$

This means that at least some vector in the set can be represented as the linear combination of other elements. Contrastingly, the set is called *linearly independent* iff

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \Rightarrow \forall i, \alpha_i = 0$$

Intuitively, if we need to find generators of a vector space, a linearly dependent set is redundant. But a linearly independent set might not be able to generate all the elements of a vector space through linear combinations. This motivates the definition of basis, which is basically a maximal linearly independent set of a vector space.

Definition 1. *Basis: A set S is called a basis of a vector space V iff S is linearly independent and any vector in V can be represented as a linear combination of elements of S.*

Notice that since any element in V can be represented as a linear combination of elements of S. This implies that adding any $v \in V \setminus S$ in S will make it linearly dependent (hence maximal linearly independent set).

One of the basic theorems of linear algebra says that the cardinality of all the basis sets is always the same and it is called the *dimension* of the vector space. Also given a linearly independent set of V, it can be extended to form a complete basis of V (keep adding linearly independent vectors at every stage).

Note: There is no mention about the uniqueness of the basis. There can be lot of basis sets for a given vector space. The span of $k < n$ elements of a basis B_1 of V (dimension n) need not be contained in the span of some k (even $n-1$) elements of B_2. Consider the standard basis $B = \{e_1, \cdots, e_n\}$ and vector $x = (1, 1, \cdots, 1)^T$. Now x or the space spanned by x is not contained in span of any $n-1$ vectors from B.

3 Inner product space

All the examples we discussed above are not just vector spaces but inner product spaces. That means they have an associated inner product. Again we won't go into the formal definition. Intuitively, inner product (dot product for $\mathbb{R}^n$) allows us to introduce the concept of angels, lengths and orthogonality between elements of vector space. We wil use $x^T y$ to denote the inner product between x and y.

Definition 2. *Orthogonality: Two elments x, y of vector space V are called* orthogonal *iff* $x^T y = 0$.

Definition 3. *Length: The length of a vector $x \in V$ is defined to be $\|x\| = \sqrt{x^T x}$.*

Using orthogonality we can come up with a simpler representation of a vector space. This requires the definition of *orthonormal basis.*

Definition 4. *A basis B of vector space V is orthonormal iff,*

- *For any two elements $x, y \in B$, $x^T y = 0$,*
- *For all elements $x \in B$, $\|x\| = 1$.*

Now every vector can be represented as a usual column vector ($n \times 1$ matrix) with respect to this orthonormal basis. It will have co-ordinates corresponding to every basis vector and operation between vectors like summation, scalar multiplication and inner product will make sense as the usual operation on the column vectors.

Given any basis of a vector space, it can be converted into an orthonormal basis. Start with a vector of the basis and normalize it (make it length 1). Take another vector, subtract the components in the direction of already chosen vectors. Normalize the remaining vector and keep repeating this process. This process always results in an orthonormal basis and is known as *Gram-Schmidt Process.*

4 Vector spaces relating to matrices

Given an $m \times n$ matrix M with real entries, we now look at few vector spaces related to matrix M.

- Image/Column space: This is the vector space spanned by columns of M. Since all columns are elements of $\mathbb{R}^m$. This vector space is the subspace of $\mathbb{R}^m$.
- Row space: This is the vector space spanned by rows of M. Since all rows are elements of $\mathbb{R}^n$. This vector space is the subspace of $\mathbb{R}^n$. It is the column space of M^T.
- Null space: It is the set of all vectors $v \in \mathbb{R}^n$, s.t., $Mv = 0$. It is the subspace of $\mathbb{R}^n$. The dimension of the null space is called the nullity of M.
- Null space of M^T: As mentioned above it is the set of vectors in $\mathbb{R}^m$, s.t., $M^T v = 0$. It is the subspace of $\mathbb{R}^m$.

Exercise 2. Verify that all of them are subspaces.

There are two basic theorems which are of importance here.

Theorem 1. *The dimension of row space is the same the dimension of column space. It is called the* rank *of M.*

This does not say that the column space and row space are the same. It only says that the dimension is same. As an exercies, give an example where row space and column space are not the same. Lets say that rank of M is r.

Theorem 2. *Rank-nullity theorem: The sum of rank and nullity of M is n.*

$$rank(M) + nullity(M) = n$$

In other words, the dimension of row space and dimension of null space for M sum up to n. Applying this to M^T, it tells us that sum of the dimension of column space and the dimension of null space of M^T is m. Actually, a stronger statement can be made, suppose $C(M), R(M), N(M), N(M^T)$ represent the four subspaces discussed above respectively. Then,

$$C(M) \oplus N(M^T) = \mathbb{R}^m \text{ and } R(M) \oplus N(M) = \mathbb{R}^n$$

This statement not just says that the dimensions add up to m/n, but that the subspaces are orthogonal and completely partition the spaces $\mathbb{R}^m/\mathbb{R}^n$.