YPH102 — Linear Vector Spaces

Lecture 2 of 3

*V.H. Belvadi**

Gram–Schmidt theorem—Schwarz inequality—Triangle inequality—Subspaces—Linear operators—Matrix elements of linear operators—Active and passive transformation

Yuvaraja's College, University of Mysore

September 2018

———

We have so far discussed what a vector space is, what vectors are, the Dirac notation and some fundamental mathematical procedures that help us handle these vectors. One of the ideas we discussed was that of orthonormal basis. We will begin this lecture by looking at a method to form such a basis.

6 Gram–Schmidt theorem

Theorem 3 (Gram–Schmidt theorem). *Given a linearly independent basis we can obtain an orthonormal basis by forming linear combinations of the given basis vectors.*

Proof. Broadly, our procedure will be as follows:

1. Scale the given vector and make it a unit vector to obtain the first basis vector.

2. Subtract the projection of the second vector along the first so we have only that part of the second vector that is perpendicular to the first.

3. Rescale by its own length the vector leftover from the previous step to obtain a second (normalised) basis vector perpendicular to the first.

And now we begin: let $|U\rangle, |V\rangle, |W\rangle, \ldots$ be a basis. The first vector of the orthonormal basis that we form using this (step 1) will be

$$|1\rangle = \frac{|U\rangle}{\sqrt{\langle U|U\rangle}}$$

To verify that the new orthonormal basis vector $|1\rangle$ is in fact normalised consider $\langle 1|1\rangle$ which is simply $\langle U|U\rangle / \sqrt{\langle U|U\rangle}^2 = 1$.

For the second vector (step 2) consider $|V\rangle$ and subtract from it that part of $|V\rangle$ which is already pointing along $|1\rangle$, i.e. $|1\rangle\langle 1|V\rangle$, so our second vector, still only orthogonal, is

$$|2'\rangle = |V\rangle - |1\rangle\langle 1|V\rangle$$

Like we proved normality for $|1\rangle$ try to prove that $|2'\rangle$ is orthogonal to $|1\rangle$ i.e. show that $\langle 1|2'\rangle = 0$.

and we proceed to also normalise it (step 3) as follows:

$$|2\rangle = \frac{|2'\rangle}{\sqrt{\langle 2'|2'\rangle}}$$

Repeat these steps for the third orthonormal basis vector:

$$|3'\rangle = |W\rangle - |1\rangle\langle 1|W\rangle - |2\rangle\langle 2|W\rangle$$

$$|3\rangle = \frac{|3'\rangle}{\sqrt{\langle 3'|3'\rangle}}$$

The procedure can be repeated for as many orthonormal basis vectors as will fit into our set. This is known as the **Gram-Schmidt theorem**. □

7 INEQUALITIES

Next we prove two exceptional theorems based on inequalities:

Theorem 4 (Schwarz inequality)**.** *The (dot/inner) product of two vectors cannot exceed the product of their lengths:*

$$|\langle U|V\rangle| \leq |U||V| \tag{9}$$

Proof. We saw in our last lecture that the inner product is positive semidefinite i.e. $\langle Z|Z\rangle \geq 0$. Let us work on this starting with the observation we have already made that $\langle Z|Z\rangle = 0 \iff |Z\rangle = 0$ as follows:

$$|Z\rangle = 0$$

$$= 1 - 1 = 1 - \frac{\langle V|V\rangle}{\sqrt{\langle V|V\rangle}^2}$$

$$= |U\rangle - \frac{\langle V|U\rangle}{|V|^2}|V\rangle$$

$$\implies \langle Z|Z\rangle = \langle U - \frac{\langle V|U\rangle}{|V|^2}V|U - \frac{\langle V|U\rangle}{|V|^2}V\rangle$$

$$= \langle U|U\rangle - \frac{\langle V|U\rangle\langle U|V\rangle}{|V|^2} - \frac{\langle V|U\rangle\langle V|U\rangle}{|V|^2} + \frac{\langle V|U\rangle\langle V|U\rangle\langle V|V\rangle}{|V|^4} \quad \geq 0$$

In the penultimate step we have opened all brackets and taken the complex

conjugate of all bra terms. Finally we get

$$\langle U|U\rangle \geq \frac{\langle V|U\rangle \langle U|V\rangle}{|V|^2}$$

$$\therefore \sqrt{\left(\sqrt{\langle V|V\rangle}\right)^2}\sqrt{\langle U|U\rangle} \geq \sqrt{\langle V|U\rangle}\sqrt{\langle U|V\rangle}$$

$$\implies |V||U| \geq |\langle V|U\rangle| \tag{9}$$

Hence proved. $\qquad\square$

Theorem 5 (Triangle inequality). *The length of a sum of vectors cannot exceed the sum of the lengths of those vectors:*

$$|U + V| \leq |U| + |V| \tag{10}$$

Proof. Start with the square of the left-hand side and use the familiar expansion of a vector U^2 into U^*U a product of the vector itself and its complex conjugate.

$$|U + V|^2 = (U + V)^*(U + V)$$

$$= U^*U + U^*V + V^*U + V^*V$$

$$= U^2 + U^*V + V^*U + V^2$$

Now we make two observations: First, for any complex number $z \equiv a + ib$ we know that $z + \bar{z} = 2a = 2\mathrm{Re}(z)$. This can be applied to the middle terms on the right-hand side above i.e. $U^*V + V^*U = U^*V + \overline{U^*V}$ giving us $2\mathrm{Re}(U^*V)$.

Second, for any complex number z we have $|z|^2 = a^2 + b^2 \geq a^2$ since b^2 is necessarily positive and removing that term will reduce the value of the right-hand side. Applying this to our U^*V term we see that $|U^*V|^2 \geq R(U^*V)$.

Combine the two observations to get

$$U^*V + V^*U = 2\mathrm{Re}(U^*V)$$

$$\leq 2|U^*V|^2 = 2|U||V|$$

Plug this into our previous result, the expansion of $|U + V|^2$, to get the triangle inequality:

$$|U + V|^2 \leq |U|^2 + 2|U||V| + |V|^2 = (|U| + |V|)^2$$

$$\implies |U + V| \leq |U| + |V| \tag{10}$$

Hence proved. (A similar 'backwards' inequality applies to the same equation but with minus signs[1].) $\qquad\square$

8 SUBSPACES

We will conclude this part of our lecture with a look at subspaces. As the name suggests these are vector spaces within vector spaces. To be precise, a **subspace** is a subset of elements of some vector space $\mathbb{V}$ which themselves form a vector space. Subsets are generally denoted by $\mathbb{V}_i^{n_i}$ where $\mathbb{V}$ is the parent space, the superscript n_i is the dimensionality of the ith subspace and the subscript i refers to the subspace itself.

An example might clear things up further: given the vector space $\mathbb{R}^2$ in the x and y directions a possible subspace could be all vectors in the $\pm x$ direction or, alternatively, in the $\pm y$ direction—they are indistinguishable. Had our parent space been $\mathbb{R}^3$ we could also have defined a subspace as the set of all vectors on the yx plane. That said, in either case, all vectors in the positive x direction alone (or negative or along the positive or negative y direction for that matter) do not form a subspace because they will not carry additive inverses. All of the examples given here carry the null vector, the origin. As for representation we label the subspace of $\mathbb{R}^3$ in, for example, the x direction as $\mathbb{R}_x^1$ and on the yx plane as $\mathbb{R}_{yz}^2$ and so on.

9 LINEAR OPERATORS AND MATRICES

9.1 *Introduction to linear operators*

Lecture 2, part 2

We now move on to the next big idea under this series of lectures: operators. An **operator** Ω is a rule or instruction that lets us transform one vector into another. It can be represented generally as follows:

$$\Omega \ket{V} = \ket{V'}$$

We call this a **transformation** by Ω from $\ket{V}$ to $\ket{V'}$. Linear operators can be scaled, i.e. $\Omega a \ket{U} = a\Omega \ket{U}$, and distributed over vectors, i.e. $\Omega\left(a \ket{U} + b \ket{V}\right) = a\Omega \ket{U} + b\Omega \ket{V}$. The same holds for bra vectors too.

The simplest linear operator is the identity operator: $I \ket{U} = \ket{U}$. Its operation is to keep a vector untouched.

Another important operator is the rotation operator $R(\theta)$ which rotates a vector by θ making it finally point along $\hat{\theta}$. Here is an example: if θ lies on the yz plane and is equal to $\pi/2$ then it rotates our $\mathbb{R}^3$ basis vectors as follows:

$R(\pi/2)\ket{1} = \ket{1}$ because $\hat{x}$ rotates around itself as $\hat{\theta}$ lies along the yz plane.

$R(\pi/2)\ket{2} = \ket{3}$ because $\hat{y}$ rotates into $\hat{z}$ along the yz plane.

$R(\pi/2)\ket{3} = -\ket{2}$ because $\hat{z}$ rotates into $-\hat{y}$ along the yz plane.

In general when we know how an operator transforms a basis we can

determine how it transforms any set of vectors in that vector space. Generally,

$$\Omega\ket{V} = \sum_i \Omega v_i \ket{i} = \sum_i v_i \Omega \ket{i} = \sum_i v_i \ket{i'}$$

Once again using $\mathbb{R}^3$ as an example if we have some vector $\ket{V}$ written as

$$\ket{V} = a\ket{1} + b\ket{2} + c\ket{3}$$

we can write down the operation of Ω on it, known how Ω operates on the basis vectors of $\mathbb{R}^3$, as follows:

$$\Omega\ket{V} = a\ket{1} + b\ket{3} - c\ket{2}$$

9.2   *Properties of linear operators*

There are a few handy properties for linear operators:

1. Their *product* simply means a sequential application of operators:

   $\Omega\Lambda\ket{U} = \Omega\left(\Lambda\ket{U}\right) = \Omega\ket{\Lambda U}.$

2. The order of operations is important as operators do not commute. Mathematically,

   $[\Omega, \Lambda] \neq 0$ where $[\Omega, \Lambda] \equiv \Omega\Lambda - \Lambda\Omega.$

3. But there are commutator identities that can prove to be of use:

   (a) $[\Omega, \Lambda\theta] = \Lambda\,[\Omega, \theta] + [\Omega, \Lambda]\,\theta$
   (b) $[\Lambda\Omega, \theta] = \Lambda\,[\Omega, \theta] + [\Lambda, \theta]\,\Omega$

4. Not all operators have inverses; those that do satisfy these two criteria:

   (a) $\Omega\Omega^{-1} = I = \Omega^{-1}\Omega$
   (b) $(\Omega\Lambda)^{-1} = \Lambda^{-1}\Omega^{-1}$

9.3   *Linear operators as matrices*

So far we have seen how a vector, in its most abstract form, can be represented in a basis by nothing more than a set of n numbers, called its components. We will now similarly represent a linear operator in a basis using n^2 numbers, specifically as an $n \times n$ square matrix.

One observation we ought to make at the outset is that, like vector components rely on a basis, matrix elements of an operator rely on a basis.

We begin by assuming we know how an operator Ω acts on the basis vectors $\ket{i}$ of some $\ket{V}$:

$$\Omega\ket{i} = \ket{i'}$$

We may write this down as a matrix Ω_{ij} using the now familiar identity

$$\braket{j | i'} = \bra{j}\Omega\ket{i} := \Omega_{ji}$$

where Ω_{ij} are n^2 elements forming the matrix of Ω *in this basis*.

Knowing this we can determine the elements of $|V\rangle$ as well. That is, we can write a matrix with elements $|v'\rangle$ as follows:

$$v'_i = \langle i|V'\rangle = \langle i|\Omega|V\rangle$$

$$= \sum_j v_j \langle i|\Omega|j\rangle \qquad \text{Since } |V\rangle = \sum_j v_j |j\rangle$$

$$= \sum_j \Omega_{ij} v_j$$

Written as a matrix this is

$$\begin{bmatrix} v'_1 \\ v'_2 \\ \vdots \\ v'_n \end{bmatrix} = \begin{bmatrix} \langle 1|\Omega|1\rangle & \langle 1|\Omega|2\rangle & \dots & \langle 1|\Omega|n\rangle \\ \langle 2|\Omega|1\rangle & & & \\ \vdots & & & \vdots \\ \langle n|\Omega|1\rangle & \dots & & \langle n|\Omega|n\rangle \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad (11)$$

On the left-hand side are the components of the transformed vector and on the extreme right-hand side are those of the original vector. In-between them, on the right-hand side, are the components of the operator written as $\langle \text{row}|\Omega|\text{column}\rangle$ where each jth column denotes the result of the action of Ω on the jth basis vector. This also holds true for bra vectors with $\langle v''| = \langle v'|\Omega$.

EXERCISE.
Find the matrix components of the $R(\pi/2)$ operator discussed earlier.

SOLUTION.
We know that $R_{ij} = \langle i|R|j\rangle$ are the elements of the matrix. We therefore have the first element as $\langle 1|R|1\rangle$ which is $\langle 1|1\rangle$ or just 1. The second element will be $\langle 1|R|2\rangle$ which is $\langle 1|3\rangle$ or zero. In a like manner we can determine all the other elements. Of particular interest is $\langle 2|R|3\rangle$ which is $\langle 2|-2\rangle$ or $-\langle 2|2\rangle$ giving us $R_{23} = -1$. In the end R turns out to be

$$R(\pi/2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

EXERCISE.
Find the matrix components of the cyclic operator that cycles $|1\rangle \rightarrow |2\rangle \rightarrow |3\rangle$.

6

10 MORE ABOUT OPERATORS

10.1 *A couple of important operators*

The **identity operator** is one we are already quite familiar with. It works as follows:

$$I_{ij} = \langle i|I|j \rangle = \langle i|j \rangle = \delta_{ij} \tag{12}$$

In other words it is a diagonal matrix with $\delta_{ii} = 1$ as its only elements.

The **projection operator** is another important operator. Consider the following expansion:

$$|V\rangle = \sum_i |i\rangle \langle i|V\rangle$$

Here we call the object $\sum_i |i\rangle \langle i|$ the projection operator and represent it as $\mathbb{P}$. If we compare this to eq. (12) and we realise that $\mathbb{P}$ is an operator that can best be described using the identity operator:

$$I = \sum_i |i\rangle \langle i| = \sum_i \mathbb{P} \tag{13}$$

where $\mathbb{P} = |i\rangle \langle i|$ is called the projection operator for $|i\rangle$ and eq. (13) is an important equation known as the **completeness relation**. It expresses the identity operator as a sum over projection operators.

Here is what the projection operator does:

$$\mathbb{P}_i |V\rangle = |i\rangle \langle i|V\rangle = |i\rangle \langle i|v_i i\rangle = |i\rangle v_i$$

Verbally, $\mathbb{P}_i$ projects out the $|i\rangle$ component of $|V\rangle$ i.e. v_i.

Finally we need to find the matrix elements of $\mathbb{P}_i$ for which we use the technique we discussed earlier for operators in general: $\Omega_{ij} = \langle i|\Omega|j \rangle$. Thus,

$$(\mathbb{P}_i)_{kl} = \langle k|i\rangle \langle i|l\rangle = \delta_{ki}\delta_{il}$$

Written as an actual matrix we get

$$\mathbb{P}_i = \begin{bmatrix} 0 & & & \cdots & & & 0 \\ & \ddots & & & & & \\ & & 0 & & & & \\ \vdots & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & & & & 0 \end{bmatrix}$$

where only the $i = k = l$ term will be 1 and the rest will be zeroes. You can also arrive at this considering $|i\rangle \langle i|$ as generic $n \times 1$ column and $1 \times n$ row matrices

with the $j = i$ element 1 and other elements 0s to get an $n \times n$ square matrix like the one shown above.

10.2   *Operator products*

Operators Ω and Λ can be multiplied and still be represented as matrices:

$$(\Omega\Lambda)_{ij} = \langle i|\Omega\Lambda|j\rangle = \langle i|\Omega I\Lambda|j\rangle = \sum_k \langle i|\Omega|k\rangle \langle k|\Lambda|j\rangle = \sum_k \Omega_{ik}\Lambda_{kj}$$

First, observe how we have introduced completeness here. Second, the matrix representing the product of operators is the product of the matrices representing their respective factors.

References

[1] Matrix analysis and applied linear algebra. C.D. Meyer: `bit.ly/2N49mao`.

[2] R. Shankar. Principles of Quantum Mechanics, 2ed. Springer.