MATHS II

Spring 2017 – Mathematics Core – Credit 4
Date: FEBRUARY 3, 2017

Instructors: B. S. Lakshmi (*Office:* A1, *E-mail:* lakshmi.burra@iiit.ac.in)
Subhadip Mitra (*Office:* A3-313C, CCNSB, *Phone:* 1587, *E-mail:* subhadip.mitra@iiit.ac.in)

TAs: Aman Rai (*E-mail:* aman.rai@students.iiit.ac.in)
Asish Varanasi (*E-mail:* varanasi.asish@students.iiit.ac.in)
Hareesh Gajulapalli (*E-mail:* hareesh.gajulapalli@students.iiit.ac.in)
Kushagra Chandak (*E-mail:* kushagra.chandak@research.iiit.ac.in)
Purnachand Jaddu (*E-mail:* purnachand.jaddu@students.iiit.ac.in)
Shikha Jain (*E-mail:* shikha.jain@students.iiit.ac.in)
Anil Rayala (*E-mail:* anil.rayala@students.iiit.ac.in)
Balaji Puppala (*E-mail:* balaji.puppala@students.iiit.ac.in)
Hemant Kasat (*E-mail:* hemant.kasat@research.iiit.ac.in)
Murali Krishna(*E-mail:* murali.dakannagari@research.iiit.ac.in)
Sanket Markan (*E-mail:* sanket.markan@students.iiit.ac.in))

Problem Set 2

Topic: Linear Vector Space

To define a **linear vector space** (LVS) (or more explicitly a LVS over the **field**[‡] of complex numbers, $\mathbb{C}$), we begin with a set S of (abstract) objects (we call them as vectors, denoted by $|\ \rangle$) and define two operations

(i) $|a\rangle + |b\rangle = |c\rangle$

(ii) $|c\rangle = \alpha\,|a\rangle$

and then set the following rules

(i) if $|a\rangle\,, |b\rangle \in S$ then $|c\rangle = |a\rangle + |b\rangle \in S$,

(ii) if $|a\rangle \in S$ and $\alpha \in \mathbb{C}$, then $\alpha\,|a\rangle \in S$ (if $\alpha \in \mathbb{R}$ then the LVS becomes a **real LVS** (RLVS) or a LVS over the field $\mathbb{R}$),

(iii) $\exists$ a null element $|0\rangle \in S$ such that $\forall\,|a\rangle \in S$, $|a\rangle + |0\rangle = |a\rangle$,

(iv) $\forall\,|a\rangle \in S$, $\exists$ an inverse (reverse) vector $|a'\rangle$ such that $|a\rangle + |a'\rangle = |0\rangle$ (we shall also use 0 in place of $|0\rangle$ since the meaning is clear),

(v) $|a\rangle + |b\rangle = |b\rangle + |a\rangle$ (commutative addition) and $(|a\rangle + |b\rangle) + |c\rangle = |a\rangle + (|b\rangle + |c\rangle)$ (associative addition),

(vi) $1\,|a\rangle = |a\rangle$,

(vii) $\alpha\,(\beta\,|a\rangle) = (\alpha\beta)\,|a\rangle$ (associative multiplication)

$(viii)$ $(\alpha+\beta)\,|a\rangle = \alpha\,|a\rangle + \beta\,|a\rangle$ and $\alpha\,(|a\rangle + |b\rangle) = \alpha\,|a\rangle + \alpha\,|b\rangle$ (distributive addition and multiplication).

Q 12. Show how

(a) the complex numbers can form a LVS,

(b) S could be the set of all complex n-column matrices (vectors) to form a LVS,

(c) the set of continuous complex valued function $f(x)$ (defined for some $a \le x \le b$) forms a LVS

(d) all polynomials in x of degree $\le n$ form a LVS and

(e) S can also be the set of all arrows lying in a plane if we interpret the multiplication with a complex number $z = re^{i\theta}$ as multiplying the length by r and rotating it by θ about the starting point of the arrow.

Q 13. We can show that any inverse vector is unique. First we assume the contrary; let $|a\rangle$ have two inverses,

$$\begin{aligned} |a\rangle + |a'\rangle &= |0\rangle \\ |a\rangle + |a''\rangle &= |0'\rangle \end{aligned}$$

where $|0'\rangle$ could be a second null vector. Now, the vector, $|a''\rangle + (|a\rangle + |a'\rangle) = |a''\rangle + |0\rangle = |a''\rangle$ (by the definition of a null vector). But, $|a''\rangle + (|a\rangle + |a'\rangle) = (|a''\rangle + |a\rangle) + |a'\rangle = (|a\rangle + |a''\rangle) + |a'\rangle = |0'\rangle + |a'\rangle = |a'\rangle + |0'\rangle = |a'\rangle$. Hence, $|a'\rangle = |a''\rangle$, i.e., the inverse is unique. Similarly, show that

(a) the null vector is also unique, and

(b) $0\,|a\rangle = |0\rangle\,, \forall\,|a\rangle \in S$. **Hint:** use 1+0=1

From two vectors we obtain a $\mathbb{C}$**-number** (a $\mathbb{R}$**-number for RLVS**) by defining a **scalar (or inner) product** as,

(i) $\langle a|b\rangle = \langle b|a\rangle^*$,

[‡] See http://mathworld.wolfram.com/Field.html for the definition of a field.

(ii) If $|d\rangle = \alpha\,|a\rangle + \beta\,|b\rangle$, then $\langle c|d\rangle = \alpha\,\langle c|a\rangle + \beta\,\langle c|b\rangle$, and

(iii) $\langle 0|0\rangle = 0$ and $\langle a|a\rangle > 0$ if $|a\rangle \neq |0\rangle$.

We shall call the two vectors, $|a\rangle$ and $|b\rangle$ **orthogonal** ($\perp$) to each other if $\langle a|b\rangle = \langle b|a\rangle = 0$. The objects $\langle a|, \langle b|, \langle c|$ etc. are **dual vectors** belonging to a different LVS. There is a one-to-one relation between vectors and dual vectors. We demand,

(i) the product of $\langle b|$ with $|a\rangle$ is identified with the scalar product $\langle b|a\rangle$, i.e., $\langle b|\cdot|a\rangle = \langle b|a\rangle$ and

(ii) the scalar product $\langle d|c\rangle$ depends linearly on $\langle d|$, i.e., $\langle d|c\rangle = \alpha^*\,\langle a|c\rangle + \beta^*\,\langle b|c\rangle$

With the scalar product defined, we can now prove the **Cauchy-Schwarz inequality** (CSI).[§] Let $|c\rangle = |a\rangle - x\,\langle b|a\rangle\,|b\rangle$ where $x \in \mathbb{R}$. Now, $\langle c|c\rangle \geq 0$ implies,

$$\begin{aligned} x^2\,\langle b|a\rangle\,\langle a|b\rangle\,\langle b|b\rangle - 2x\,\langle b|a\rangle\,\langle a|b\rangle + \langle a|a\rangle &\geq 0 \qquad (2.1)\\ \Rightarrow\quad (1 - x\langle b|b\rangle)^2 + \left(\frac{\langle a|a\rangle\,\langle b|b\rangle}{|\langle a|b\rangle|^2} - 1\right) &\geq 0 \end{aligned}$$

Since this is valid for any real x,

$$\begin{aligned} \langle a|a\rangle\,\langle b|b\rangle &\geq \langle a|b\rangle\,\langle b|a\rangle = |\langle a|b\rangle|^2 \\ \sqrt{\langle a|a\rangle}\sqrt{\langle b|b\rangle} &\geq |\langle a|b\rangle|\,. \qquad (2.2) \end{aligned}$$

Our proof is not valid if $\langle a|b\rangle = 0$, but in this case the CSI is trivially satisfied (the same is true if $|a\rangle$ or $|b\rangle$ is a null vector).

Remember a quadratic $ax^2 + 2bx + c$ (where $a,b,c,x \in \mathbb{R}$) ≥ 0 if $a > 0$ and the discriminant $4(b^2 - acb) \leq 0$. (When the discriminant is negative, the quadratic has no real roots, i.e., it never crosses the zero line for any value of x and when the discriminant is zero, the quadratic has only one real root, i.e., it touches the zero line from one side for only one value of x, the extremum. The condition $a > 0$ ensures that the extremum is a minimum.) This can also be used to prove the CSI from Eq. (2.1).

The familiar notion of distance defined in the usual 3D real LVS of position vectors is generalized as a **metric**. A set S is called a metric space if $\forall\ a,b \in S$, a real non-negative number $\rho(a,b)$ can be associated with the pair such that,

(i) $\rho(a,b) = \rho(b,a)$

(ii) $\rho(a,b) = 0 \quad \Rightarrow \quad a = b$

(iii) $\rho(a,b) + \rho(b,c) \geq \rho(a,c)$ **(triangle inequality)**.

A LVS with a scalar product is a metric space. Like, we can define

$$\rho^2(a,b) = (\langle a| - \langle b|)\cdot(|a\rangle - |b\rangle) = (\langle a|a\rangle + \langle b|b\rangle) - (\langle a|b\rangle + \langle b|a\rangle)\,. \qquad (2.3)$$

Q 14. Show that the null vector is orthogonal to every vector in a LVS and it is the only such vector.

Q 15. Show that $\rho(a,b)$ defined in Eq. (2.3) satisfies all the defining properties of a metric.

Q 16. Check

(a) that in our familiar 3D real LVS of position vectors, the dot product $\langle A|B\rangle = \vec{A}\cdot\vec{B} = |\vec{A}||\vec{B}|\cos\theta_{AB}$ is a scalar product and the CSI holds. However, show that $\langle A|B\rangle' = a_1 b_1 + \frac{1}{2}a_2 b_2 + \frac{1}{3}a_3 b_3$ is also a valid scalar product (a_i and b_i's are the components of $|A\rangle$ and $|B\rangle$ along the x-y-z axes) and the CSI holds too. This implies that the scalar product is not unique. (Note, however, the second definition of the scalar product is not frame independent. Suppose the component of $|A\rangle$ are $(1,0,0)$ in one frame. Now, if we rotate to a new frame whose y-axis coincides with the present x-axis then see how $\langle A|A\rangle'$ changes.)

(b) that in the LVS of complex numbers, the definition $\langle z_1|z_2\rangle = z_1^* z_2$ works as a scalar product and the CSI holds. Define $\rho(z_1, z_2)$ following Eq. (2.3) and show explicitly that it is a metric. Will $\langle z_1|z_2\rangle = \mathrm{Re}(z_1^* z_2)$ work as a scalar product?

(c) whether we can form a RLVS with complex numbers. Can we set $\langle z_1|z_2\rangle = \mathrm{Re}(z_1^* z_2)$ as the scalar product for this RLVS?

(d) that in the LVS of complex n-column matrices, the definition $\langle Y|X\rangle = Y^\dagger X$ works as a scalar product and the CSI holds.

(e) that the vector space of 2×2 complex matrices form a LVS. Then show that $\langle Y|X\rangle = \frac{1}{2}\mathrm{Tr}\left(Y^\dagger X\right)$ defines a scalar product in this space.

Q 17. Show that

[§] en.wikipedia.org/wiki/Augustin-Louis_Cauchy & en.wikipedia.org/wiki/Hermann_Schwarz

(a) the set of all continuous complex valued functions of $x \in \mathbb{R}$ defined over $0 \le x \le 1$ form a LVS. Show that

$$\langle g|f\rangle = \int_0^1 g^*(x) f(x) dx$$

defines a valid scalar product. Verify the CSI by considering two vectors $f(x) = x^2 - 1$ and $g(x) = 5 + 2ix$.

(b) $P_1(x) = x$ and $P_3(x) = \left(5x^3 - 3x\right)/2$ are orthogonal vectors.

- Consider a set of N vectors from a LVS: $|1\rangle, |2\rangle, |3\rangle, \ldots, |N\rangle$. They are called **linearly independent** (LI) if

$$\sum_{i=1}^{N} a_i |i\rangle = 0 \quad \Rightarrow \quad a_i = 0,\ \forall\, i = \{1, 2, 3, \ldots, N\}, \tag{2.4}$$

otherwise they are **linearly dependent** (LD). If N non-null vectors are LD, then at least one of them can be expressed as a linear combination of the others.

- The **dimensionality of a LVS** is equal to the maximum number of LI vectors that can be found in it. Suppose we start with a non-null vector, and then we start identifying (or constructing) other vectors so that the collection satisfies Eq. (2.4) (i.e., they are LI of each other). If N is the maximum number of such vectors that collectively satisfy Eq. (2.4), then we say the LVS is N dimensional. If the process of finding such vectors never ends then the LVS is **infinite dimensional**.
- In an N-dimensional LVS, finding a set of N-LI vectors means all other vectors from the LVS can be expressed as linear combinations of these vectors or, in other words, any set of N-LI vectors constitutes a **basis**. Given a basis, decomposition of a given vector in terms of this basis is unique.
- Calculations often become simple if one chooses an **orthonormal basis**, a set of N vectors $\{|e_i\rangle; i = 1, 2, 3, \ldots, N\}$ for which $\langle e_i|e_j\rangle = \delta_{ij}$ (i.e. the unit vectors). Note that a set of non-null vectors, $|x_i\rangle$ that are orthogonal ($\perp$) to each other, i.e., $\langle x_i|x_j\rangle = 0$ if $i \ne j$, are also LI. In an orthonormal basis, $|a\rangle = \sum_{i=1}^{N} a_i |i\rangle$ with $a_i = \langle e_i|a\rangle$.
- Given a set of N-LI vectors, it is always possible to choose an orthonormal basis. A process known as the **Gram-Schmidt orthonormalization**¶ lets us find such a basis. Suppose we have a set of N-LI vectors, $\{|x_i\rangle; i = 1, 2, 3, \ldots, N\}$. We start by defining a unit vector along $|x_1\rangle$ as,

$$|e_1\rangle = \frac{1}{\sqrt{\langle x_1|x_1\rangle}} |x_1\rangle. \tag{2.5}$$

We then take $|x_2\rangle$ and define the second unit vector $|e_2\rangle$ which is $\perp |e_1\rangle$ by taking away the component of $|x_2\rangle$ along $|e_1\rangle$,

$$|e_2\rangle = \frac{1}{\sqrt{\left(\langle x_2| - \langle x_2|e_1\rangle\langle e_1|\right)\cdot\left(|x_2\rangle - \langle e_1|x_2\rangle|e_1\rangle\right)}}\left(|x_2\rangle - \langle e_1|x_2\rangle|e_1\rangle\right). \tag{2.6}$$

By construction they have unit magnitudes, i.e., $\langle e_1|e_1\rangle = \langle e_2|e_2\rangle = 1$ and it is easy to see that $\langle e_1|e_2\rangle = \langle e_2|e_1\rangle = 0$. We can continue this process iteratively. Suppose we have already constructed $|e_i\rangle$, the i^{th} unit vector out of the vectors, $|x_1\rangle, \ldots, |x_i\rangle$, we can now construct the $(i+1)^{\text{th}}$ unit vector by constructing a vector out of $|x_{i+1}\rangle$ that is orthogonal to all the unit vectors constructed so far, (i.e., $|e_1\rangle, \ldots, |e_i\rangle$) and then normalizing it:

$$\begin{aligned} |e_{i+1}\rangle &= \frac{1}{N_{i+1}}\left(|x_{i+i}\rangle - \sum_{k=1}^{i}\langle e_k|x_{i+1}\rangle|e_k\rangle\right), \\ \text{where,}\quad N_{i+1} &= \sqrt{\left(\langle x_{i+i}| - \sum_{k=1}^{i}\langle x_{i+1}|e_k\rangle\langle e_k|\right)\cdot\left(|x_{i+i}\rangle - \sum_{k=1}^{i}\langle e_k|x_{i+1}\rangle|e_k\rangle\right)}. \end{aligned} \tag{2.7}$$

Q 18. Prove that if N non-null vectors are LD then at least one of them can be expressed as a linear combination of the others.

Q 19. Prove that a set of non-null vectors, $|x_i\rangle$ that are orthogonal ($\perp$) to each other, i.e., $\langle x_i|x_j\rangle = 0$ if $i \ne j$, are also LI.

Q 20. Prove that decomposition of a given vector in terms of a given basis is unique.

Q 21. Prove that in an N-dimensional LVS, any set of N-LI vectors constitutes a basis, i.e., all vectors can be expressed as linear combinations of these vectors.

Q 22. In the real 3D LVS of position vectors you are given three vectors, $|x_1\rangle = \hat{\imath} + 2\hat{\jmath}$, $|x_2\rangle = -\hat{\jmath} + \hat{k}$ and $|x_3\rangle = \hat{\imath} - \hat{\jmath} + \hat{k}$. Confirm that these three are LI and then follow the Gram-Schmidt orthonormalization to construct an orthonormal basis and explicitly verify that the vectors are orthonormal.

¶ en.wikipedia.org/wiki/Jørgen_Pedersen_Gram & en.wikipedia.org/wiki/Erhard_Schmidt