Lecture 6b: Inner Product Spaces

Suresh Govindarajan

Department of Physics, Indian Institute of Technology Madras

April 2020

Length of Vectors – Two Examples

- We begin with the usual vectors i.e., elements of $\mathbb{R}^3$. Given a vector $\mathbf{u} = (u_1, u_2, u_3)^T$, we associate a length, usually denoted by $|\mathbf{u}|$, to it.

  $$|\mathbf{u}|^2 = (u_1)^2 + (u_2)^2 + (u_3)^2 \geq 0 \ .$$

  A closlely related object is the dot/scalar product of two vectors. Let $\mathbf{v} = (v_1, v_2, v_3)^T$. Then

  $$\mathbf{u} \cdot \mathbf{v} = u_1\, v_1 + u_2\, v_2 + u_3\, v_3 \ .$$

- The second example is the one of the polarization vector i.e. an element $\boldsymbol{\mathcal{E}} = (\mathcal{E}_1, \mathcal{E}_2)^T \in \mathbb{C}^2$. The length of the this vector is (note the need for complex conjugation)

  $$|\boldsymbol{\mathcal{E}}|^2 = \mathcal{E}_1^*\, \mathcal{E}_1 + \mathcal{E}_2^*\, \mathcal{E}_2 \geq 0 \ .$$

  Recall that the time-averaged energy density was $\frac{1}{2\epsilon_0}$ times the above length.

Inner Product

The inner product generalizes the notion of dot product of vectors to arbitrary linear vector spaces. Several of its properties follow from the two examples that we just saw. Let $\mathbf{u}, \mathbf{v} \in \mathbb{V}$, a linear vector space over $\mathbb{F}$.

- The inner product of these two vectors is written as

$$\langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{F} \ .$$

  It takes as input two vectors and returns a scalar like the scalar product of two vectors in $\mathbb{R}^3$.

- We call $\langle \mathbf{u}, \mathbf{u} \rangle$ the square of the norm of the vector $\mathbf{u}$. One writes

$$||\mathbf{u}||^2 = \langle \mathbf{u}, \mathbf{u} \rangle \in \mathbb{R}_{\geq 0} \ .$$

We will state its properties next and one can easily verify that the scalar product satisfies all the properties.

Properties of an Inner Product

Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{V}$, a linear vector space over $\mathbb{C}$ and $a, b \in \mathbb{C}$.

1. Conjugation Symmetry

$$\langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle^*$$

2. Linearity in the second argument

$$\langle \mathbf{u}, a\,\mathbf{v} + b\,\mathbf{w} \rangle = a\,\langle \mathbf{u}, \mathbf{v} \rangle + b\,\langle \mathbf{u}, \mathbf{w} \rangle$$

   Caution: In most maths books, inner products are defined to be linear in their first argument. We follow a convention that is followed in quantum mechanics.

3. Positivity of norm

$$\langle \mathbf{u}, \mathbf{u} \rangle \geq 0\ ,$$

   and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ implies $\mathbf{u} = \mathbf{0}$.

Remarks on Inner Products

- Condition 1 for a real LVS implies that the inner product is symmetric in its arguments.
- Condition 1 implies that $\langle \mathbf{u}, \mathbf{u} \rangle \in \mathbb{R}$. This is weaker than condition 3 which says it is positive.
- Conditions 1 and 2 imply

$$\langle a\,\mathbf{v} + b\,\mathbf{w}, \mathbf{u} \rangle = a^* \, \langle \mathbf{v}, \mathbf{u} \rangle + b^* \, \langle \mathbf{w}, \mathbf{u} \rangle$$

  The inner product is anti-linear (or conjugate linear) in its first argument.

Example: Let $\mathbb{V} = \mathbb{R}^n$ or $\mathbb{C}^n$ and let $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)^T$. Then $(\mathbf{u}^\dagger := (\mathbf{u}^*)^T = (u_1^*, u_2^*, \ldots, u_n^*))$

$$\langle \mathbf{u}, \mathbf{v} \rangle := \mathbf{u}^\dagger \mathbf{v} = \sum_{i=1}^{n} (u_i)^* \, v_i \ ,$$

is an inner product on $\mathbb{C}^n$. Check this!

- A linear vector space with an inner product is called an inner product space.
- Two vectors are said to be orthogonal if their inner product is zero.
- The Cauchy-Schwarz inequality holds. (Proof in PS 13)

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq ||\mathbf{u}|| \; ||\mathbf{v}|| \; .$$

  In $\mathbb{R}^3$, this reduces to $|\cos\theta| \leq 1$ where θ is the angle between the two vectors **u** and **v**.
- The triangle inequality holds: (Proof in PS 13)

$$||\mathbf{u}|| + ||\mathbf{v}|| \geq ||\mathbf{u} + \mathbf{v}||$$

- To every vector **v**, the inner product gives us a linear function that we will denote by $\langle \mathbf{v}, \quad \rangle$

$$\langle \mathbf{v}, \quad \rangle \text{ maps } \quad \mathbf{u} \rightarrow \langle \mathbf{v}, \mathbf{u} \rangle \in \mathbb{F} \quad \text{ for all } \mathbf{u} \in \mathbb{V} \; .$$

Orthonomal basis

Recall that without an inner product, linear independence was the only condition imposed on a basis set. In an inner product space, we have extra structure. In analogy with $\mathbb{R}^3$, we can impose two extra conditions on a basis $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ for an inner product space $\mathbb{V}$ of dimension n.

- Orthogonality $\langle e_i, e_j \rangle = 0$ for all $i \neq j$ and
- Normalization $\langle e_i, e_i \rangle = 1$ for all $i = 1, \ldots, n$.

Equivalently, $\langle e_i, e_j \rangle = \delta_{ij}$. We call such a basis an orthonormal (ON) basis. Given a basis, not necessarily ON, there is a iterative procedure to construct an ON basis. This is called the Gram-Schmidt orthonormalization process. You may have already seen this in your Maths course.

Gram-Schmidt Orthonormalization

We begin with a basis $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ and after n steps construct an orthonormal basis $\widehat{\mathcal{B}} = (\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n)$.

Step 1: Set $\hat{e}_1 = \frac{e_1}{||e_1||}$. $\hat{e}_1$ is normalied to unity.

Step 2: Let $e_2' = e_2 - \langle \hat{e}_1, e_2 \rangle \hat{e}_1$. By construction, e_2' is orthogonal to $\hat{e}_1$ i.e., $\langle \hat{e}_1, e_2' \rangle = 0$. Check!
Define $\hat{e}_2 = \frac{e_2'}{||e_2'||}$ to normalise it to unity.

Step i Let $i > 2$. Define $e_i' = e_i - \sum_{j=1}^{i-1} \langle \hat{e}_j, e_i \rangle \hat{e}_j$. By construction, e_i' is orthogonal to $\hat{e}_j$ for all $j < i$. Define $\hat{e}_i = \frac{e_i'}{||e_i'||}$ to normalise it to unity. At the end of this step, we have a set of orthonormal vectors $(\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_i)$ i.e., any pair of distinct vectors are orthogonal and each of them have unit norm.

At the end of the n-th step, we obtain the required ON basis, $\widehat{\mathcal{B}} = (\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n)$.

An example

Consider the following basis for $\mathbb{R}^3$.

$$\mathcal{B} = \{e_1, e_2, e_3\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\} .$$

Step 1 $\hat{e}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Step 2 $e_2' = \begin{pmatrix} -2/3 \\ 1/3 \\ 1/3 \end{pmatrix}$ and $\hat{e}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$.

Step 3 $e_3' = \begin{pmatrix} 0 \\ -1/2 \\ 1/2 \end{pmatrix}$ and $\hat{e}_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$.

We thus obtain the ON basis.

$$\widehat{\mathcal{B}} = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$$

Can we use this as an ON basis for $\mathbb{C}^3$?

Dirac Notation

Paul A M Dirac, who won the 1933 Nobel Prize in Physics, contributed immensely to the development of quantum mechanics and quantum electrodynamics. He introduced a notation that is used to this day in quantum mechanics.

- Vectors $\mathbf{u}, \mathbf{v} \in \mathbb{V}$ are written as $|\mathbf{u}\rangle, |\mathbf{v}\rangle$ – this is read as "ket u, ket v" and so on.
- The linear maps $\langle \mathbf{u}, \quad \rangle, \langle \mathbf{v}, \quad \rangle$ are written as $\langle \mathbf{u}|, \langle \mathbf{v}|$ – this is read as "bra u, bra v" and so on.
- The inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ is written as $\langle \mathbf{u}|\mathbf{v} \rangle$ – just replace the comma with the bar.

Dirac invented the names bra and ket as a breaking up of a "bra(c)ket" which is what he called the inner product.

Change of basis

Let $\mathcal{B} = (|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle)$ and $\mathcal{B}' = (|e'_1\rangle, |e'_2\rangle, \ldots, |e'_n\rangle)$ be another ON basis for an inner product space $\mathbb{V}$. ON implies

$$\langle e_i | e_j \rangle = \langle e'_i | e'_j \rangle = \delta_{ij} \ .$$

Recall that these two basis are related by an invertible matrix S

$$|e_i\rangle = \sum_{j=1}^{n} |e'_j\rangle \ S_{ji} \qquad (*)$$

Taking the inner product of $(*)$ with $|e'_k\rangle$, we get

$$\langle e'_k | e_i \rangle = \sum_{j=1}^{n} \langle e'_k | e'_j \rangle \ S_{ji} = \sum_{j=1}^{n} \delta_{kj} \ S_{ji} = S_{ki} \ .$$

Taking the inner product of $(*)$ with $|e_k\rangle$, we get

$$\delta_{ki} = \langle e_k | e_i \rangle = \sum_{j=1}^{n} \langle e_k | e'_j \rangle \ S_{ji} = \sum_{j=1}^{n} (S^\dagger)_{kj} S_{ji} = (S^\dagger \cdot S)_{ki} \ .$$

Unitary matrices

Let us collect the two results that we obtain in the previous slide: The change of basis matrix is

$$S_{ki} = \langle e'_k | e_i \rangle \implies S^*_{ki} = \langle e_i | e'_k \rangle =: S^\dagger_{ik}$$

where the matrix $S^\dagger$ to be the conjugate transpose matrix of S. The second condition then becomes

$$\boxed{S^\dagger \cdot S = I_n}\,, \qquad \text{or} \qquad S^\dagger = S^{-1}$$

where I_n is the $n \times n$ identity matrix. Matrices that satisfy the above condition are called unitary matrices. Changes of basis that map one ON basis to another are thus called a unitary change of basis.

Properties of unitary matrices: 1. $|\det(S)| = 1$.

2. The product of two unitary matrices is also unitary.

Linear Operators

Let T be a linear operator, i..e, a map from $\mathbb{V}$ to itself. For an inner product space with ON basis $\mathcal{B}$, there is a simple expression for the matrix of T: (Recall $T(e_j) = \sum_k e_k T_{kj}$)

$$T_{ij} = \langle e_i | T(e_j) \rangle := \langle e_i | T | e_j \rangle \ .$$

The operator $T^\dagger$ is defined as follows

$$\langle \mathbf{u} | T(\mathbf{v}) \rangle^* = \langle T(\mathbf{v}) | \mathbf{u} \rangle =: \langle \mathbf{v} | T^\dagger(\mathbf{u}) \rangle \qquad \text{for all } \mathbf{u}, \mathbf{v} \in \mathbb{V} \ .$$

The matrix for $T^\dagger$ is the conjugate transpose of the matrix for T. (Caution: In maths books, the symbol T^* is used.)
Definition: A linear operator T is hermitian if $T = T^\dagger$.

Eigenvalues of hermitian matrices are always real. Such operators play an important role in quantum mechanics.

Examples of matrices

- The most general 2×2 unitary matrix takes the form

$$U = \begin{pmatrix} a & b \\ -b^* e^{i\varphi} & a^* e^{i\varphi} \end{pmatrix} \quad a, b \in \mathbb{C} \text{ and } \varphi \in [0, 2\pi) \ ,$$

  with $|a|^2 + |b|^2 = 1$. Check that this matrix is unitary.

- The most general hermitian 2×2 matrix is a real linear combination of the the identity matrix and the following three hermitian matrices called the Pauli matrices.

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad ,$$

$$\text{and} \quad \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ .$$