Linear Vector Space

Hsiu-Hau Lin
hsiuhau@phys.nthu.edu.tw
(Mar 29, 2010)

The notes cover linear vector space and special matrices (Boas 3.9-3.10).

• Index notation

As mentioned in previous notes, the matrix multiplication can be expressed in index notation,

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}. \tag{1}$$

Familiarity of the index notation can help to prove several useful theorems. For instance, the identity matrix can be expressed in terms of the Kronecker delta,

$$\delta_{ij} = \begin{cases} 1, & \text{for } i = j, \\ 0, & \text{for } i \neq j, \end{cases} \tag{2}$$

It then becomes clear that any matrix multiplied by the identity matrix remains the same, $\sum_k \delta_{ik} M_{kj} = M_{ij}$. The Kronecker delta is useful for other purpose as well. For instance, the following integral is of essential importance for Fourier series and can be expressed elegantly by the Kronecker delta,

$$\int_{-\pi}^{\pi} \cos mx \cos nx \, dx = \pi \delta_{mn}.$$

Making use of the index notation, we can prove that the transpose of the product of several matrices equals the product of the transpose matrices in opposite order,

$$(ABCD)^T = D^T C^T B^T A^T. \tag{3}$$

Let's take the product of two matrices as an illustrating example,

$$[(AB)^T]_{ij} = (AB)_{ji} = \sum_k A_{jk} B_{ki} = \sum_k (A^T)_{kj} (B^T)_{ik} = (B^T A^T)_{ij}.$$

The rule is simple – when taking the transpose of the product, one just need to take the transpose for each matrix and multiply them in opposite order.

Another interesting theorem is about the cyclic invariance of the trace,

$$\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB). \tag{4}$$

The proof is again straightforward in the index notation,

$$\mathrm{Tr}(ABC) = \sum_{ijk} A_{ij} B_{jk} C_{ki} = \sum_{ijk} C_{ki} A_{ij} B_{jk} = \mathrm{Tr}(CAB).$$

It is important to emphasize that $\mathrm{Tr}(ABC)$ and $\mathrm{Tr}(BAC)$ do not equal to each other in general. Let us consider an interesting counter example for the trace theorem,

$$A = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{2} & 0 & \dots \\ 0 & 0 & 0 & \sqrt{3} & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad A^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ \sqrt{1} & 0 & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

These two matrices are annihilation and creation operators for the simple harmonic oscillator in quantum mechanics. Compute the product of these two matrices,

$$AA^\dagger = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ 0 & 2 & 0 & 0 & \dots \\ 0 & 0 & 3 & 0 & \dots \\ 0 & 0 & 0 & 4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad A^\dagger A = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & 3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The difference between the products $AA^\dagger, A^\dagger A$ is just the identity matrix,

$$AA^\dagger - A^\dagger A = \mathbf{1}. \tag{5}$$

However, strange thing happens when taking the trace for the above relation. For the left-hand side, the trace should be zero according to the theorem but the trace of the right-hand side is infinite,

$$\mathrm{Tr}(AA^\dagger) - \mathrm{Tr}(A^\dagger A) = 0 \neq \mathrm{Tr}\ \mathbf{1}!$$

Since infinity is clearly not equal to zero, we must make some mistake somewhere. Can you figure out where the discrepancy lies?

• Linear vector space

It is common practice that we can describe the position of a particle by a three dimensional vector. Similarly, the velocity needs another three dimensional vectors. For a system consists of N particles, we thus need $2N$ vectors in the three dimensional space to capture its motion. But, we can also construct a $6N$ dimensional vector in an abstract space to describe the motion of the system. This $6N$ dimensional space has a name – phase space. You will learn more about the phase space in classical mechanics.

One can also use vector notation to label different genome sequences,

$$\begin{aligned} \boldsymbol{A} &= (0,0,1,1,0,0,1,1,0,0), \\ \boldsymbol{B} &= (1,1,1,0,1,0,0,0,1,0). \end{aligned}$$

Between two different genome sequences, the notion of "distance" is still meaningful and often referred as the Hamming distance,

$$d(\boldsymbol{A},\boldsymbol{B}) = \sum_{i=1}^{n}(A_i - B_i)^2 = \sum_{i=1}^{n}(1-\delta_{A_iB_i}) = n - \sum_{i=1}^{n}\delta_{A_iB_i}.$$

The sequence space is quite compact and crowded. As you can readily tell, each sequence has as many neighbors as the number of base pairs n but the maximal distance between any two sequences is only n. Because of the huge number of neighbors and the close distance, it is also nicknamed as Manhattan measure.

• Four-dimensional spacetime

In special relativity, the motion of a particle is better described by a four-dimensional vector. The well-known Lorentz transformation can be casted into matrix form,

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma u/c & 0 & 0 \\ -\gamma u/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}, \tag{6}$$

where $\gamma = 1/\sqrt{1-(u/c)^2}$. When the velocity is slow (in comparison to the speed of light), $\gamma \approx 1$ and the Lorentz transformation simplify into the Galilean transformation. Since the transformation for the transverse directions are trivial, we can shrink the 4×4 matrix into 2×2 and concentrate on the non-trivial coordinates. As can be verify by straightforward algebra, the matrix elements can be expressed by hyperbolic functions,

$$\cosh\alpha = \gamma = \frac{1}{\sqrt{1-(u/c)^2}}, \qquad \sinh\alpha = \gamma(u/c) = \frac{u/c}{\sqrt{1-(u/c)^2}}. \tag{7}$$

The Lorentz transformation in the reduced dimensions (t, x) is

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\alpha & -\sinh\alpha \\ -\sinh\alpha & \cosh\alpha \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}. \tag{8}$$

The above matrix resembles the rotation matrix in two dimensions if the hyperbolic functions are replaced by the sinusoidal ones. There is a deeper connection between them because of similar group structure.

The matrix form is not just for "good looking" but can be helpful from time to time. For instance, the rule of matrix multiplication provides a simple route to derive the velocity addition rule in special relativity. For a particle moving at velocity v, its motion satisfies $x = vt$. In the moving frame, its motions becomes $x' = v't'$ and we would like to figure out the relation between v and v'. Starting from the relation $x = vt$, the Lorentz transformation gives

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\alpha(ct) - \sinh\alpha(vt) \\ -\sinh\alpha(ct) + \cosh\alpha(vt) \end{pmatrix}.$$

The velocity v' observed in the moving frame can be found,

$$v' = \frac{x'}{t'} = \frac{(v/c)\cosh\alpha - \sinh\alpha}{\cosh\alpha - (v/c)\sinh\alpha} c = \frac{v-u}{1-uv/c^2}.$$

Note that, if $v = c$ (motion of light) in the rest frame, $v' = c$ in the moving frame and echoes the assumption that the speed of light is invariant.

• Orthonormal basis: Gram-Schmidt method

Schwarz inequality for n-dimensional Euclidean space is

$$\left(\sum_{i=1}^{n} A_i B_i\right)^2 \leq \left(\sum_{i=1}^{n} A_i^2\right)\left(\sum_{i=1}^{n} B_i^2\right), \quad \rightarrow \quad |\boldsymbol{A}\cdot\boldsymbol{B}| \leq AB. \tag{9}$$

The proof involves a central idea of projection. Choosing the vector

$$\boldsymbol{C} = B\boldsymbol{A} - (\boldsymbol{A}\cdot\boldsymbol{B})\boldsymbol{B}/B,$$

compute its length and the inequality follows. The projection can help to construct orthogonal vectors. Take two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ as example,

$$\boldsymbol{B}' = \boldsymbol{B} - (\boldsymbol{B}\cdot\hat{\boldsymbol{a}})\hat{\boldsymbol{a}} \quad \rightarrow \quad \boldsymbol{B}'\cdot\boldsymbol{A} = 0.$$

Generalizing the above idea, one can build up an orthonormal basis by Gram-Schmidt method. This method is best learnt by working through examples. Consider the following three vectors,

$$\boldsymbol{A} = (0,0,5,0), \quad \boldsymbol{B} = (2,0,3,0), \quad \boldsymbol{C} = (7,1,-5,3).$$

The first unit vector in the orthonormal basis can be constructed by normalizing the first vector,

$$\hat{\boldsymbol{e}}_1 \;=\; \boldsymbol{A}/A = (0,0,1,0).$$

To construct the second unit vector, one needs to subtract the projection along the first unit vector,

$$\boldsymbol{B} - (\boldsymbol{B}\cdot\hat{\boldsymbol{e}}_1)\hat{\boldsymbol{e}}_1 = (2,0,0,0), \quad \rightarrow \quad \hat{\boldsymbol{e}}_2 = (1,0,0,0).$$

Following the same spirit, the last one can be found by subtracting projections along the previous directions,

$$\boldsymbol{C} - (\boldsymbol{C}\cdot\hat{\boldsymbol{e}}_1)\hat{\boldsymbol{e}}_1 - (\boldsymbol{C}\cdot\hat{\boldsymbol{e}}_2)\hat{\boldsymbol{e}}_2 = (0,1,0,3), \quad \rightarrow \quad \hat{\boldsymbol{e}}_3 = \frac{1}{\sqrt{10}}(0,1,0,3).$$

The generalization to higher dimensions is straightforward but the algebra can be tedious. Luckily, we have computers these days. You only need to be familiar with the algorithm of the Gram-Schmidt method – write a program and leave the messy algebra to your computer.

• Complex vector space: Hermitian and unitary matrices

As you may have guessed, there is no reason to limit our imagination – vectors can be complex as well. The inner product between two vectors is not necessarily real anymore. Consider two complex vectors

$$\boldsymbol{A} = \begin{pmatrix} 3i \\ 1-i \\ 2+3i \\ 1+2i \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} -1 \\ 1+2i \\ 3-i \\ i \end{pmatrix},$$

there are two types of inner products (complex conjugate to each other),

$$\boldsymbol{A}^\dagger \boldsymbol{B} = 4 - 4i = \boldsymbol{B}^\dagger \boldsymbol{A}.$$

In quantum mechanics, a particle is described by a complex vector named wave function. Two types of matrices are often encountered in physics – Hermitian and unitary matrices. The angular momentum along the z-axis is captured by a Hermitian matrix,

$$L_z = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \text{ and its square is } \quad L_z^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{1}.$$

As explained in class before, the rotation matrix can be obtained by exponentiating the angular momentum,

$$\begin{aligned} e^{-iL_z\theta} &= 1 + (-iL_z\theta) + \frac{(-iL_z\theta)^2}{2!} + \frac{(-iL_z\theta)^3}{3!} + \frac{(-iL_z\theta)^4}{4!} + \frac{(iL_z\theta)^5}{5!} + ... \\ &= \mathbf{1}(\cos\theta) - iL_z(\sin\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \end{aligned}$$

This relation is not a coincidence. Construct a matrix U by exponentiating a Hermitian matrix H,

$$U = e^{iH}, \qquad U^\dagger = e^{-iH^\dagger} = e^{-iH}, \tag{10}$$

one can show that the matrix U is unitary because $e^{iH}e^{-iH} = 1$.