附录: 线性代数概要

陈童

September 3, 2020

这个数学附录的内容是为了帮助初学者理解线性代数的基本概念和思维方式，并不是要给出完备的严格数学证明。因此下面讲述的很多内容在数学严格性上肯定是有待补充的。

1 向量空间

通常的三维空间$\mathbb{R}^3$由如下有序三元实数组构成

$$\mathbb{R}^3 = \{[x, y, z] : x, y, z \in \mathbb{R}\}.$$

我们可以记$\mathbf{x} = [x, y, z]$，称之为$\mathbb{R}^3$空间的向量，任取$\mathbf{x}_1 = [x_1, y_1, z_1]$, $\mathbf{x}_2 = [x_2, y_2, z_2]$，我们可以定义向量的加法，$\mathbf{x}_1 + \mathbf{x}_2 = [x_1 + x_2, y_1 + y_2, z_1 + z_2]$。显然，这样定义的向量加法满足加法交换律和结合律。同时我们还可以定义标量乘法$\lambda\mathbf{x} = [\lambda x, \lambda y, \lambda z], \lambda \in \mathbb{R}$。显然，它满足分配律，即$\lambda(\mathbf{x}_1 + \mathbf{x}_2) = \lambda\mathbf{x}_1 + \lambda\mathbf{x}_2$。

类似的代数结构我们也可以在n维复空间$\mathbb{C}^n$中发现，$\mathbb{C}^n$由如下有序n元复数组构成

$$\mathbb{C}^n = \{[x_1, x_2, ..., x_n] : x_1, x_2, ..., x_n \in \mathbb{C}\}.$$

记$\mathbf{x} = [x_1, x_2, ..., x_n]$, $\mathbf{y} = [y_1, y_2, ..., y_n]$, 称为$\mathbb{C}^n$中的两个向量。类似地可以定义向量加法，$\mathbf{x} + \mathbf{y} = [x_1 + y_1, x_2 + y_2, ..., x_n + y_n]$, 很显然，它同样满足交换律和结合律。同样可以定义标量乘法，$\lambda\mathbf{x} = [\lambda x_1, \lambda x_2, ..., \lambda x_n]$，不过现在标量$\lambda \in \mathbb{C}$。很明显，这样的标量乘法也满足分配律。

利用向量的加法和标量乘法，我们很容易将k个向量线性组合起来，得到一个新的向量，

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + ... + c_k\mathbf{x}_k.$$

式中向量$\mathbf{x}$就是$\mathbf{x}_1$,$\mathbf{x}_2$,...,$\mathbf{x}_k$的线性组合。可以进行任意的线性组合就是向量的核心特征。反过来，具有这个特征的东西都叫向量。

比方说我们考察$[0,1]$区间上的连续复数值函数，我们把所有这些函数的集合记为$C_{[0,1]}$。任给$f(x), g(x) \in C_{[0,1]}$，很显然我们可以定义一个新的$C_{[0,1]}$集合内的函数$(f+g)(x)$, $(f+g)(x) = f(x) + g(x)$。这就给$C_{[0,1]}$集合内的函数定义了加法，它显然满足交换律和结合律。我们也可以定义标量乘法, $(\lambda f)(x) = \lambda f(x)$，它显然满足分配律。有了这两种运算，我们就可以把$C_{[0,1]}$内的函数任意线性组合，比方说$c_1 f_1(x) + c_2 f_2(x) + ... + c_k f_k(x)$显然还是$C_{[0,1]}$内的函数，$c_1, c_2, ..., c_k$就是组合系数，按照标量乘法的定义，它们当然都是普通的复数。可见，我们完全可以把$C_{[0,1]}$内的函数称作向量！

把这些要素抽象出来，我们就可以给出向量空间的一般性定义：*向量空间V是一个集合，它里面的元素(我们称作向量)可以是任何东西，只要这些东西之间定义了加法和标量乘法运算，并且这些运算满足交换律、结合律以及分配律。简言之，只要这些东西可以进行任意的线性组合，组合系数是普通的数。当然，还有很重要的一条，无论如何线性组合，结果依然得是集合V中的一个向量，也即是说，集合V在加法和标量乘法运算下得具有封闭性。*

向量和向量空间是抽象的。然而读者通常都能按照3维空间里面的我们从高中就已经熟悉的普通向量来想象它们，因为不管向量空间多么抽象，它和3维向量空间都有同样的本质，那就是里面的向量可以任意线性组合。所以，通常人们也把**向量空间称作线性空间**。

既然向量是抽象的，我们最好给它们发明一个通用的抽象的记号。我们把向量记作$|\psi\rangle$，读作向量ψ。比方说在前面的$\mathbb{R}^3$空间情形，$|\psi\rangle$就是三元组$[x, y, z]$，所以$|\psi\rangle = [x, y, z]$。在$\mathbb{C}^n$情形，$|\psi\rangle = [x_1, x_2, ..., x_n]$。而在$C_{[0,1]}$情形，$|\psi\rangle = f(x)$。

在量子力学中我们关心的都是组合系数为复数的向量空间，称为复向量空间。根据向量空间的定义，对于任何复向量空间V，假如向

量$|\psi_1\rangle, |\psi_2\rangle, ..., |\psi_k\rangle \in V$，则它们的任意复系数线性组合也必属于$V$，即

$$c_1|\psi_1\rangle + c_2|\psi_2\rangle + ... + c_k|\psi_k\rangle \in V.$$

子空间与直和

如果向量空间V的一个子集U本身也是向量空间，即U本身就在加法和标量乘法下保持封闭，则U就称为V的一个**子空间**。比如很容易验证$\{[x_1, x_2, 0] : x_1, x_2 \in \mathbb{C}\}$就是$\mathbb{C}^3$的一个子空间。同样，通常3维空间$\mathbb{R}^3$中任何一个过原点的平面也构成$\mathbb{R}^3$的一个子空间。实际上，人们通常都可以把向量空间的子空间想象成“一个过原点的超平面”。

设U和W均是向量空间V的子空间，并且如果对于$|u\rangle \in U$, $|w\rangle \in W$，当且仅当$|u\rangle = |w\rangle = 0$时才满足$|u\rangle + |w\rangle = 0$。则我们就可以定义$U$和$W$的**直和**, 记为$U \oplus W$，它也是$V$的子空间，$U \oplus W$的要点在于其里面的任何向量$|\psi\rangle$均可以唯一分解成如下形式

$$U \oplus W = \{|\psi\rangle = |u\rangle + |w\rangle : |u\rangle \in U, |w\rangle \in W\}.$$

这种分解的唯一性很容易证明，设另有$|u'\rangle \in U$, $|w'\rangle \in W$也满足$|\psi\rangle = |u'\rangle + |w'\rangle$, 则将$|\psi\rangle$的这两种分解方式相减就可以得到$(|u'\rangle - |u\rangle) + (|w'\rangle - |w\rangle) = 0$，根据前面的假设必有$|u'\rangle - |u\rangle = 0$, $|w'\rangle - |w\rangle = 0$，这就证明了分解的唯一性。

举个例子，假设向量空间$\mathbb{R}^3$的两个子空间U和W的定义如下，$U = \{[x, y, 0] : x, y \in \mathbb{R}\}$, $W = \{[0, 0, z] : z \in \mathbb{R}\}$, 则$U \oplus W = \mathbb{R}^3$。

直和的概念当然可以推广。设$U_1, U_2, ..., U_m$均是向量空间V的子空间，并且对于$|u_1\rangle \in U_1, ..., |u_m\rangle \in U_m$，当且仅当$|u_1\rangle = |u_2\rangle = ... = |u_m\rangle = 0$时才有$|u_1\rangle + |u_2\rangle + ... + |u_m\rangle = 0$。那么我们就可以定义这$m$个子空间的直和，记为$U_1 \oplus U_2 \oplus ... \oplus U_m$。具体来说，$U_1 \oplus ... \oplus U_m$也是$V$的一个子空间，它里面的向量可以唯一分解成如下形式

$$U_1 \oplus ... \oplus U_m = \{|u_1\rangle + |u_2\rangle + ... + |u_m\rangle : |u_1\rangle \in U_1, ..., |u_m\rangle \in U_m\}.$$

举个例子。设U_j是$\mathbb{C}^n$中除第j个坐标以外其余坐标全是0的那些向量所组成的子空间，则$U_1 \oplus U_2 \oplus ... \oplus U_n = \mathbb{C}^n$。

2 有限维向量空间

张成空间与线性无关

复向量空间V中的一组向量$|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$的所有可能线性组合所构成的集合称为$|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$的张成空间，记为$\mathrm{Span}(|v_1\rangle, |v_2\rangle, ..., |v_m\rangle)$, 即

$$\mathrm{Span}(|v_1\rangle, |v_2\rangle, ..., |v_m\rangle) = \{c_1|v_1\rangle + c_2|v_2\rangle + ... + c_m|v_m\rangle : c_1, c_2, ..., c_m \in \mathbb{C}\}.$$

可以证明$\mathrm{Span}(|v_1\rangle, |v_2\rangle, ..., |v_m\rangle)$是包含$|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$的最小子空间。

有限维向量空间就是可以由向量空间中一组有限数目的向量张成的向量空间。除了有限维向量空间之外的向量空间都叫无限维的。虽然量子力学中的向量空间常常为无限维，但其基本数学结构都可以从有限维的情况得到理解。因此，我们这个附录将主要限于讨论有限维向量空间。而且，我们将主要讨论有限维复向量空间。

举个例子。很明显，由n个向量组成的向量组$|e_1\rangle = [1, 0, ..., 0], |e_2\rangle = [0, 1, 0, ..., 0], ..., |e_n\rangle = [0, ..., 0, 1]$张成了向量空间$\mathbb{C}^n$。

再举个例子，任何两个不超过n次的复多项式的和依然是一个不超过n次的复多项式，所以所有不超过n次的复多项式$p(z)$的集合构成了一个复向量空间，定义这个向量空间的向量$|p\rangle = p(z)$。很显然，这个向量空间可以由如下向量组张成，$|0\rangle = 1, |1\rangle = z^1, ..., |n\rangle = z^n$。这是因为，任何不超过$n$次的复多项式都可以写成$p(z) = c_0 + c_1 z + c_2 z^2 + ... + c_n z^n$，用向量的形式来写即是$|p\rangle = c_0|0\rangle + c_1|1\rangle + c_2|2\rangle + ... + c_n|n\rangle$。

对于一组向量$|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$，如果$c_1|v_1\rangle + c_2|v_2\rangle + ... + c_m|v_m\rangle = 0$, 当且仅当所有的系数都为0，即$c_1 = c_2 = ... = c_m = 0$, 则我们就称$|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$**线性无关**。否则就称它们线性相关。容易证明，如果$|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$线性无关，则任何$|u\rangle \in \mathrm{Span}(|v_1\rangle, |v_2\rangle, ..., |v_m\rangle)$ 的线性组合形式$|u\rangle = a_1|v_1\rangle + a_2|v_2\rangle + ... + a_m|v_m\rangle$都必定是**唯**一的。因为否则，假设$|u\rangle$另有一种不同的线性组合$|u\rangle = b_1|v_1\rangle + b_2|v_2\rangle + ... + b_m|v_m\rangle$, 则将$|u\rangle$的两种线性组合相减就会得到$(a_1 - b_1)|v_1\rangle + (a_2 - b_2)|v_2\rangle + ... + (a_m - b_m)|v_m\rangle = 0$, 其中组合系数不全为0，但这就和$|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$线性无关的假设矛盾了。

如果$|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$线性相关，那这就说明对于张成子空间$W = \mathrm{Span}(|v_1\rangle, |v_2\rangle, ..., |v_m\rangle)$来说，向量组$|v_1\rangle, |v_2\rangle, ..., |v_m\rangle$有冗余, 我们完全可

以从中删去某些向量，使得剩余的向量依然能够张成W。比方说，向量组$|e_1\rangle = [1,0], |e_2\rangle = [0,1], |e_3\rangle = [1,1]$张成了向量空间$\mathbb{C}^2$，但是由于$|e_1\rangle + |e_2\rangle - |e_3\rangle = 0$，因此这个向量组是线性相关的，我们完全可以从中删除向量$|e_3\rangle$，剩下的$|e_1\rangle, |e_2\rangle$依然张成了向量空间$\mathbb{C}^2$。

所以对于张成一个固定的子空间来说，线性无关向量组是最小的向量组。由此我们可以得到下面的概念。

基和维数

向量空间V的基是任何一组既线性无关又能张成V的向量，基中的每一个向量叫作基向量。比方说，对于所有不超过n次的复多项式所成的向量空间，$|0\rangle = 1, |1\rangle = z^1, ..., |n\rangle = z^n$就是它的一组基。再比方说，对于向量空间$\mathbb{C}^n$, $|e_1\rangle = [1,0,...,0], |e_2\rangle = [0,1,0,...,0], ..., |e_n\rangle = [0,...,0,1]$就是它的一组基。

根据基的定义我们知道，给定向量空间的一组基，则空间中任何向量都可以写成这组基的线性组合形式，并且组合系数是唯一的。在有限维向量空间V上选定一组基$|e_1\rangle, |e_2\rangle, ..., |e_n\rangle$以后，$V$中的任意向量$|v\rangle$就可以唯一性地用它的组合系数来表示。具体来说，我们可以唯一性地将$|v\rangle$写成，

$$|v\rangle = v_1|e_1\rangle + v_2|e_2\rangle ... + v_n|e_n\rangle, \tag{1}$$

式中$v_1, v_2, ..., v_n \in \mathbb{C}$为复组合系数。进一步将(1)式重写成

$$|v\rangle = (|e_1\rangle, |e_2\rangle, ..., |e_n\rangle) \begin{pmatrix} v_1 \\ v_2 \\ ... \\ v_n \end{pmatrix}, \tag{2}$$

等式右边的表达式表示将行与列的各个分量对应地乘起来再求和，结果刚好由(1)式给出。从(2)式我们很容易看出，向量$|v\rangle$可以唯一性地表示成列矢量$v = \left(\begin{smallmatrix} v_1 \\ v_2 \\ ... \\ v_n \end{smallmatrix}\right)$。人们也很容易验证，向量的加法就刚好对应相应列矢量的加法，向量的标量乘法就对应列矢量的标量乘法。所以任何向量空间在取定基以后都可以表示成列矢量空间。正是在这个意义上，我们完全可以将向量$|v\rangle$想象成一个列矢量。

但是，一个向量空间的基本身不是唯一的。比方说，对于向量空间$\mathbb{C}^n$。$|e_1\rangle = [1, 0, ..., 0], |e_2\rangle = [0, 1, 0, ..., 0], ..., |e_n\rangle = [0, ..., 0, 1]$是它的一组基，但是将这组基中的第$n$个向量$|e_n\rangle$ 替换成$|e'_n\rangle = [0, ..., 0, 1, 1]$结果依然是一组基。

然而由于基是最小的张成向量组，所以对于有限维向量空间来说，基矢量的数目是唯一的。有限维向量空间V的基向量的数目就称为V的维数，记作$\dim(V)$。比方说，$\dim(\mathbb{C}^n) = n$。

设V为有限维向量空间，再设它能分解成子空间$U_1, U_2, ..., U_m$的直和，即$V = U_1 \oplus ... \oplus U_m$, 则可以证明

$$\dim(V) = \dim(U_1) + \dim(U_2) + ... + \dim(U_m).$$

3 线性算符

向量空间V上的线性算符T就是V到自身的一个映射，$T : V \to V$。对于V的任意向量$|u\rangle$，T将会把它映射到V的一个新向量$|v\rangle$，记作$|v\rangle = T|u\rangle$，并将$T|u\rangle$读作算符T在向量$|u\rangle$上的作用。并且，这种作用得保持向量空间V的线性结构，即保持向量空间的加法和标量乘法。具体来说就是，对于任意$|u\rangle, |w\rangle \in V$, 必有

$$T(|u\rangle + |w\rangle) = T|u\rangle + T|w\rangle.$$

另外，对于任意$\lambda \in \mathbb{C}$, $|u\rangle \in V$，有

$$T(\lambda|u\rangle) = \lambda(T|u\rangle).$$

比方说，对于所有复多项式构成的向量空间，如下定义的运算T是一个线性算符，$Tp(z) = zp(z)$, 式中$p(z)$为任意一个多项式，T的作用是将这个多项式乘以z，结果显然还是一个多项式。同样，下面的求导运算D也是复多项式向量空间上的线性算符，$Dp(z) = \frac{dp(z)}{dz}$。

再比方说，容易验证下式定义的T是向量空间$\mathbb{C}^2$上的线性算符, $T([x, y]) = [2x - y, 7x + 5y]; x, y \in \mathbb{C}$。

有一个特殊的线性算符值得单独解释一下，那就是恒等算符，也叫做单位算符，我们常常将它简记为1，它在任何向量上的作用结果依然是这个向量本身，即$1|u\rangle = |u\rangle$。

我们可以按照下面的式子分别定义两个线性算符A, B的和$A+B$，以及它们的乘积AB，

$$
\begin{aligned}
(A+B)|u\rangle &= A|u\rangle + B|u\rangle, \\
(AB)|u\rangle &= A(B|u\rangle).
\end{aligned}
$$

读者容易验证，$A+B$和AB都依然是线性算符。

给定线性算符T, 如果向量$|u\rangle$满足$T|u\rangle = 0$，则称之为算符T的零向量。很明显，如果$|u\rangle, |v\rangle$均为T的零向量，则$|u\rangle + |v\rangle$也为T的零向量，而$\lambda|u\rangle$这样的$|u\rangle$乘上标量常数λ以后的向量也是T的零向量。也即是说，T的所有零向量所构成的空间$\mathrm{Null}(T)$在加法和标量乘法下封闭。因此$\mathrm{Null}(T)$构成整个向量空间的一个子空间，称为算符T的零空间。

比方说，由于常数的导数为0，前面我们定义的作用在所有复多项式的向量空间上的线性算符D的零空间就是0次多项式空间，即$\mathrm{Null}(D) = \{a : a \in \mathbb{C}\}$。

对于线性算符A，若存在某个算符B，使得$AB = BA = 1$，则我们说，A, B互为逆算符，记作$A = B^{-1}$, $B = A^{-1}$。

如果算符A有逆算符，我们就说它可逆。很容易证明，A可逆的必要条件是$\mathrm{Null}(A) = \{0\}$。这是因为，如果A可逆，则我们就可以用A^{-1}去乘零向量的方程$A|u\rangle = 0$，从而得到$|u\rangle = 0$。实际上可以进一步证明，如果A是有限维向量空间上的线性算符，则$\mathrm{Null}(A) = \{0\}$其实是A可逆的充要条件。

线性算符与矩阵

在n维向量空间V上取定一组基$|e_1\rangle, |e_2\rangle, ..., |e_n\rangle$，注意到$V$中的任何向量都可以用基向量来展开，进而设线性算符$A$在基向量上的作用结果可以用基展开为

$$
A|e_j\rangle = \sum_{i=1}^{n} |e_i\rangle a_{ij}, \quad j = 1, 2, .., n \tag{3}
$$

式中a_{ij}为复组合系数，请注意式中a_{ij}下标的顺序。人们常常将所有a_{ij}排成

一个矩阵，记为$\widehat{A}$,

$$\widehat{A}=\begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & ... \\ ... & ... & ... & ... \\ a_{n1} & ... & ... & a_{nn} \end{pmatrix}.$$

由此我们很容易注意到(3)式$A|e_j\rangle$的结果为，行$(|e_1\rangle, |e_2\rangle, ..., |e_n\rangle)$与矩阵$\widehat{A}$的第$j$列各分量对应相乘并求和。进而我们就可以将(3)式重新写成

$$\begin{aligned} A(|e_1\rangle, |e_2\rangle, ..., |e_n\rangle) &= (A|e_1\rangle, A|e_2\rangle, ..., A|e_n\rangle) \\ &= (|e_1\rangle, |e_2\rangle, ..., |e_n\rangle)\begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & ... \\ ... & ... & ... & ... \\ a_{n1} & ... & ... & a_{nn} \end{pmatrix}. \end{aligned}$$

也即是说，取定了基以后，线性算符A可以唯一性地表示成矩阵$\widehat{A}$。

下面我们再来考察A在向量$|v\rangle$上的作用$A|v\rangle = |u\rangle$可以如何表示。为此我们将$|v\rangle$在基中展开为$|v\rangle = \sum_{j=1}^n v_j|e_j\rangle$，注意到$A|v\rangle = A(\sum_{j=1}^n v_j|e_j\rangle) = \sum_{j=1}^n v_j A|e_j\rangle = \sum_{j,i=1}^n |e_i\rangle a_{ij}v_j = \sum_{i=1}^n(\sum_{j=1}^n a_{ij}v_j)|e_i\rangle$。显然最后结果中的$(\sum_{j=1}^n a_{ij}v_j)$为矩阵$\widehat{A}$的第$i$行各分量与列矢量$v = \left(\begin{smallmatrix} v_1 \\ v_2 \\ ... \\ v_n \end{smallmatrix}\right)$各分量对应相乘并求和。记$u_i = \sum_{j=1}^n a_{ij}v_j$, 则有$A|v\rangle = \sum_{i=1}^n u_i|e_i\rangle = |u\rangle$, 其中

$$\begin{pmatrix} u_1 \\ u_2 \\ ... \\ u_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & ... \\ ... & ... & ... & ... \\ a_{n1} & ... & ... & a_{nn} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ ... \\ v_n \end{pmatrix}.$$

很显然，最终的这个结果说明，取定了基以后，我们可以将向量表示成列矢量，将线性算符表示成矩阵，而算符在向量上的作用就正好对应于矩阵与列矢量相乘！也正是在这个意义上，我们可以粗略地将抽象的向量想象成列矢量，将抽象的线性算符想象成矩阵。

不仅如此，我们还可以证明，取定了基以后，线性算符的和可以表示成相应矩阵的和，而线性算符的乘积可以表示成相应矩阵的矩阵乘积。

以算符乘积为例。假设有线性算符A和B, 其中A在基向量上的作用如前所述，而B在基向量上的作用为$B|e_j\rangle = \sum_{i=1}^n |e_i\rangle b_{ij}$，其中$b_{ij}$为复矩阵$\widehat{B}$的第$i$行第$j$列。则$AB|e_k\rangle = A(B|e_k\rangle) = A(\sum_{j=1}^n |e_j\rangle b_{jk}) = \sum_{j=1}^n A|e_j\rangle b_{jk} = \sum_{i,j=1}^n |e_i\rangle a_{ij} b_{jk} = \sum_i^n |e_i\rangle (\sum_{j=1}^n a_{ij} b_{jk})$。进而根据算符与矩阵之间的对应关系我们可以知道，$AB$乘积算符的矩阵$\widehat{AB}$的第$i$行第$k$列必为

$$(\widehat{AB})_{ik} = \sum_{j=1}^n a_{ij} b_{jk}. \tag{4}$$

这个方程告诉我们，矩阵$\widehat{AB}$的第i行第k列由矩阵$\widehat{A}$的第i行各分量与矩阵$\widehat{B}$的第k列各分量对应相乘并求和。通常人们把这个矩阵乘法规则归纳为如下方程

$$\widehat{AB} = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & ... \\ ... & ... & ... & ... \\ a_{n1} & ... & ... & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & ... & b_{1n} \\ b_{21} & b_{22} & ... & ... \\ ... & ... & ... & ... \\ b_{n1} & ... & ... & b_{nn} \end{pmatrix} = \widehat{A}\widehat{B}.$$

由此可见，算符的乘积正好表示成了相应矩阵的矩阵乘积。正是在这个意义上，我们有时候会说，线性算符和矩阵本质上是一回事。

4 不变子空间、本征值与本征向量

为了更好地理解向量空间V上的线性算符A, 我们常常将V分解为若干个真子空间的直和，

$$V = U_1 \oplus U_2 \oplus ... \oplus U_m. \tag{5}$$

然后我们可以将算符A限制在某个真子空间U_j上来考虑，记为$A|_{U_j}$，它表示仅仅将A作用在子空间U_j内的向量上。这样做是因为，限制以后的$A|_{U_j}$往往会比原来的整个算符A更简单，更好理解。

但是，$A|_{U_j}$不一定是U_j上的线性算符，也就是说，虽然$A|_{U_j}$作用在U_j上，但它不一定将U_j映射到U_j本身，有可能有某些U_j里的向量会被映射到U_j外面去！如果是这种情形，那单独研究$A|_{U_j}$其实就没有任何简化。所以为了能真正简化问题，我们通常要求上面分解出来的每一个U_j在$A|_{U_j}$的作用

下都封闭，即任何一个U_j中的向量在A的作用下都依然映射到U_j中的某个向量。或者说，$A|_{U_j}$将把U_j映射到U_j。这时候，我们就称这些子空间为算符A的**不变子空间**。可见，不变子空间的本质就是要在算符A的作用之下保持封闭。

例如$\mathrm{Null}(A)$就是A的一个不变子空间。这是因为，首先零向量0显然属于$\mathrm{Null}(A)$。其次，根据$\mathrm{Null}(A)$的定义，对于任意$|u\rangle \in \mathrm{Null}(A)$，必有$A|u\rangle = 0 \in \mathrm{Null}(A)$，从而$\mathrm{Null}(A)$在$A$的作用下的确封闭。

对于某个非零向量$|v\rangle$，如果一维子空间$U_v = \{c|v\rangle : c \in \mathbb{C}\}$是算符$A$的不变子空间，那么就称向量$|v\rangle$为算符$A$的**本征向量**。根据不变子空间的定义，这时候必有$A|v\rangle \in U_v$，从而也即是说，$A|v\rangle$必定具有$\lambda|v\rangle$的形式，即

$$A|v\rangle = \lambda|v\rangle.$$

这个方程就称为算符A的本征方程，其中的λ就称为A的**本征值**。

当然我们也可以在向量空间上选定一组基，进而把一切都翻译成矩阵。比方说，取基以后，我们可以将向量$|v\rangle$表示成列矢量v，把算符A表示成矩阵$\widehat{A}$，那上面的本征方程就可以表示成一个矩阵方程，$\widehat{A}v = \lambda v$，也即是

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix}.$$

人们也称这个方程为矩阵$\widehat{A}$的本征方程。

我们可以把算符A的本征方程改写成$(A - \lambda \cdot 1)|v\rangle = 0$, 其中$|v\rangle$为非零向量。可见，$\lambda$为算符$A$本征值的充要条件是，$\mathrm{Null}(A - \lambda \cdot 1)$中包含有非零向量。对于有限维向量空间，根据算符可逆的充要条件，这等价于$A - \lambda \cdot 1$不可逆。所以，对于有限维向量空间，**λ为算符A的本征值等价于$A - \lambda \cdot 1$不可逆**。

但是，给定一个本征值λ，可能有多个线性无关的向量满足同样的本征方程。也即是说，方程$(A - \lambda \cdot 1)|v\rangle = 0$可能有多个线性无关的本征向量解。不过很显然，所有这些本征向量解的全体就构成了$\mathrm{Null}(A - \lambda \cdot 1)$，人们也常常称之为$A$的本征值为$\lambda$的本征空间，记为$E(\lambda, A)$。因此，

$$E(\lambda, A) = \mathrm{Null}(A - \lambda \cdot 1).$$

5 内积空间

内积

对于高中学过的三维向量空间$\mathbb{R}^3$中的向量$\mathbf{x}=[x_1,x_2,x_3]$，我们可以定义它的模长$|\mathbf{x}|$为$|\mathbf{x}|^2=x_1^2+x_2^2+x_3^2$。如果另有一个向量$\mathbf{y}=[y_1,y_2,y_3]$，那我们还可以定义两个向量的内积$\mathbf{x}\cdot\mathbf{y}$为$\mathbf{x}\cdot\mathbf{y}=x_1y_1+x_2y_2+x_3y_3$。

推广到n维实向量空间$\mathbb{R}^n$中的两个向量$\mathbf{x}=[x_1,x_2,...,x_n]$和$\mathbf{y}=[y_1,y_2,...,y_n]$，我们可以将它们的内积定义为$\mathbf{x}\cdot\mathbf{y}=x_1y_1+x_2y_2+...x_ny_n$。特别的，$\mathbf{x}$的模长$|\mathbf{x}|$定义为$|\mathbf{x}|^2=\mathbf{x}\cdot\mathbf{x}$。

但是，在量子力学中，我们主要关心的是复向量空间。对于最简单的复向量空间$\mathbb{C}^n$中的向量$\mathbf{z}=[z_1,z_2,...,z_n]$和$\mathbf{w}=[w_1,w_2,...,w_n]$ 我们需要把内积的概念推广成，$\overline{w}_1z_1+\overline{w}_2z_2+...+\overline{w}_nz_n$，式中$\overline{w}$表示复数$w$的复共轭，有时候也写成$w^*$。为什么要在内积的定义里面多加一个复数共轭我们很快就会清楚。但很显然，这样定义以后，作内积的两个向量就不完全对称了。我们可以利用前面发明的抽象向量记号$|z\rangle=\mathbf{z},|w\rangle=\mathbf{w}$来反映这种不对称性，具体来说，我们将$|z\rangle$和$|w\rangle$的内积记为$\langle w|z\rangle$，它由下式计算

$$\langle w|z\rangle=\overline{w}_1z_1+\overline{w}_2z_2+...+\overline{w}_nz_n.$$

注意这个公式，我们需要对内积符号左边的那个向量进行复数共轭。

之所以将左边的向量加上一个复共轭，原因在于向量的模长由向量与其自身的内积确定。而只有这样定义以后，向量$|z\rangle$的模长公式$\big||z\rangle\big|^2=\langle z|z\rangle$才是一个大于零的正实数(因为复数$z$只有和它的复共轭$\overline{z}$相乘结果才是正实数)。

将以上的例子作一个一般性的抽象，我们就可以在任何复向量空间V上定义内积运算。所谓的内积，就是一个函数，它把V中的任何一对向量$|u\rangle,|v\rangle$映射到一个复数$\langle u|v\rangle$。并且这个映射满足：(1). **正定性**，即对于任何$|v\rangle\in V$，$\langle v|v\rangle\geq 0$, 等于号仅当$|v\rangle=0$时才成立。(2). **线性性**，即$\langle u|v\rangle$对于符号右边的向量来说是线性的。具体来说即是满足$\langle u|\big(c_1|v_1\rangle+c_2|v_2\rangle\big)=c_1\langle u|v_1\rangle+c_2\langle u|v_2\rangle$。(3). **共轭对称性**，即$\langle u|v\rangle^*=\langle v|u\rangle$。定义了内积的复向量空间就称为内积空间。

对于高中学过的三维向量，我们知道，两个向量的内积为零又称作这两个向量正交。这个概念同样可以推广到任何内积空间。如果$\langle u|v\rangle=0$, 我们

就称$|u\rangle$和$|v\rangle$**正交**。类似的，类比于单位向量的概念，如果一个向量$|u\rangle$模长为1，即满足$\langle u|u\rangle = 1$，我们就说它是归一的。

正交归一基与向量表示

对于一个有限维内积空间V, 如果它的一组基$|e_1\rangle, |e_2\rangle, ..., |e_n\rangle$满足

$$\langle e_i|e_j\rangle = \delta_{ij}, i, j = 1, 2, ..., n.$$

我们就称这一组基为正交归一基。这里克龙内克符号δ_{ij}当两个指标相同时取1，不同时取0。

取定了正交归一基以后，我们就可以把向量$|v\rangle$展开成，$|v\rangle = \sum_{i=1}^{n} v_i|e_i\rangle$。通过利用正交归一关系计算内积$\langle e_i|v\rangle$我们容易发现，

$$v_i = \langle e_i|v\rangle.$$

由此我们也可以得到$v_i^* = \langle e_i|v\rangle^* = \langle v|e_i\rangle$(利用了内积的共轭对称性)。

现在我们来考察$\langle u|v\rangle$, 我们有$\langle u|v\rangle = \langle u|\sum_{i=1}^{n} v_i|e_i\rangle = \sum_{i=1}^{n} v_i\langle u|e_i\rangle$, 假设把$|u\rangle$展开成$|u\rangle = \sum_{i=1}^{n} u_i|e_i\rangle$，则利用$u_i^* = \langle u|e_i\rangle$，我们可以得到

$$\langle u|v\rangle = \sum_{i=1}^{n} u_i^* v_i. \tag{6}$$

前面我们已经知道，向量$|u\rangle, |v\rangle$在基中可以分别表示为列矢量u, v, 它们由下式给出

$$u = \begin{pmatrix} u_1 \\ u_2 \\ ... \\ u_n \end{pmatrix} = \begin{pmatrix} \langle e_1|u\rangle \\ \langle e_2|u\rangle \\ ... \\ \langle e_n|u\rangle \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \\ ... \\ v_n \end{pmatrix} = \begin{pmatrix} \langle e_1|v\rangle \\ \langle e_2|v\rangle \\ ... \\ \langle e_n|v\rangle \end{pmatrix}.$$

则很显然，上面的(6)式其实就是

$$\langle u|v\rangle = \begin{pmatrix} u_1^* & u_2^* & ... & u_n^* \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ ... \\ v_n \end{pmatrix} = u^\dagger v. \tag{7}$$

式中$u^\dagger$表示列矢量u的复共轭再转置(称之为厄密共轭)，它显然是行矢量。(7)式告诉我们，可以将内积符号左边的向量想象成一个行矢量，它是相应列矢量的厄密共轭。

基于以上观察，我们干脆可以定义一种左矢$\langle u|$，其定义就是

$$\langle u| = |u\rangle^\dagger.$$

左矢可以想象成行矢量，是列矢量的厄密共轭。因此，内积运算最终的公式形式(7)就是行矢量乘以列矢量。注意，行矢量在左边，列矢量在右边。

正交补空间

最后我们再给出一个简单的概念。设V为内积空间，U为V的一个子空间。则U的正交补空间也是一个子空间，记为$U^\perp$, 它的定义是

$$U^\perp = \{|v\rangle \in V : 对每个|u\rangle \in U均有\langle u|v\rangle = 0\}.$$

也即是说，$U^\perp$由V中所有与U正交的向量构成。从而，整个向量空间V可以正交分解成U和它的正交补$U^\perp$的直和，即

$$V = U \oplus U^\perp.$$

关于内积空间以及内积空间中的线性算符的更多讨论，由于和量子力学有直接的关系，所以我们放在本书的正文中进行。