科普文之直和与直乘

小李哥
Institute of Physics, Chinese Academy of Sciences

摘要

最近计算狄拉克方程、泡利矩阵、分析乱七八糟拓扑分类搞得的心烦，遂在此把相关矩阵运算公式整理一下。虽然这都是大一的内容，但舍己为人是吾辈之使命。

1 直和

给定两个向量空间 V, W，维数分别是 n, m。V 里有一组向量基底 $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$，同样 W 里也有一组基底 $\left\{\vec{f}_1, \vec{f}_2, \cdots, \vec{f}_m\right\}$. 我们现在定义 $n+m$ 个向量构成基组 $\left\{\vec{e}_1, \cdots, \vec{e}_n, \vec{f}_1, \cdots, \vec{f}_m\right\}$. 由这些基组生成的线性空间叫做 V 与 W 生成的直和空间 $V \oplus W$，线性空间维数为 $n+m$.
任意向量 $\vec{v} \in V$ 都可以写成基底线性叠加的形式 $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + \cdots + v_n\vec{e}_n$. 线性代数里我们可以用这些叠加系数构成的列向量来代表这个向量。

$$\vec{v} = \left.\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}\right\} n \tag{1}$$

值得注意，这个列向量是线性空间 V 的一个元素，列向量有 n 个分量。特别地，向量基底可以写成

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}_V, \quad \cdots, \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}_V \tag{2}$$

若基组是正交组，$\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$，则有 $v_i = \vec{e}_i \cdot \vec{v}$. 同样对于 w 中的任意向量也可以有

$$\vec{w} = \left.\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}\right\} m \tag{3}$$

基向量为

$$\vec{f}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_W, \quad \vec{f}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}_W, \quad \cdots, \quad \vec{f}_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}_W \tag{4}$$

自然地，对于元素 $\vec{v}\in V$ 和 $\vec{w}\in W$ 在直和空间的列向量形式可以推广成

$$\vec{v}=\left.\begin{pmatrix} v_1\\ \vdots\\ v_n\\ \hline 0\\ \vdots\\ 0\end{pmatrix}\right\} n+m,\quad \vec{w}=\left.\begin{pmatrix} 0\\ \vdots\\ 0\\ w_1\\ \vdots\\ w_m\end{pmatrix}\right\} n+m \tag{5}$$

可以看到，只是把其余部分用 0 填充。

我们也可以定义两个非零向量的直和

$$\vec{v}\oplus\vec{w}=\left.\begin{pmatrix} v_1\\ \vdots\\ v_n\\ \hline w_1\\ \vdots\\ w_m\end{pmatrix}=\begin{pmatrix}\vec{v}\\ \vec{w}\end{pmatrix}\right\} n+m \tag{6}$$

数学上矩阵表示一个线性空间到另一个线性空间的映射。这里我们只讨论一个 $V\mapsto V$ 的特殊情况，因为量子力学里大部分情况便是如此。力学量 $A:V\mapsto V$

$$\vec{v}\mapsto A\vec{v} \tag{7}$$

$$\begin{pmatrix} v_1\\ v_2\\ \vdots\\ v_n\end{pmatrix}\longmapsto\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{22} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}_V\begin{pmatrix} v_1\\ v_2\\ \vdots\\ v_n\end{pmatrix}_V \tag{8}$$

类似的有矩阵 $B:W\mapsto W,\vec{w}\mapsto B\vec{w}$

在直和空间里，我们希望矩阵 A,B 经过某种运算后作用在向量上保持 $\vec{v}\mapsto A\vec{v}$ 和 $\vec{w}\mapsto B\vec{w}$。这样的运算记成 $A\oplus B$. 运算结果可以利用分块对角矩阵构造

$$A\oplus B=\left(\begin{array}{c|c} A & 0_{n\times m}\\ \hline 0_{m\times n} & B\end{array}\right) \tag{9}$$

例如 $n=2,m=3$ 的情况

$$A=\begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}_V,\quad B=\begin{pmatrix} b_{11} & b_{12} & b_{13}\\ b_{21} & b_{22} & b_{23}\\ b_{31} & b_{32} & b_{33}\end{pmatrix}_W \tag{10}$$

他们的直和为

$$A\oplus B=\left(\begin{array}{cc|ccc} a_{11} & a_{12} & 0 & 0 & 0\\ a_{21} & a_{22} & 0 & 0 & 0\\ \hline 0 & 0 & b_{11} & b_{12} & b_{13}\\ 0 & 0 & b_{21} & b_{22} & b_{23}\\ 0 & 0 & b_{31} & b_{32} & b_{33}\end{array}\right) \tag{11}$$

显然 A,B 分别作用在不同的子空间里。显然直和矩阵 $A\oplus B$ 作用在直和向量 $\vec{v}\oplus\vec{w}$ 为

$$(A\oplus B)(\vec{v}\oplus\vec{w})=\left(\begin{array}{c|c} A & 0_{n\times m}\\ \hline 0_{m\times n} & B\end{array}\right)\begin{pmatrix}\vec{v}\\ \vec{w}\end{pmatrix}=\begin{pmatrix}A\vec{v}\\ B\vec{w}\end{pmatrix}=(A\vec{v})\oplus(B\vec{w}) \tag{12}$$

若有两个直和矩阵相乘，利用对角矩阵运算显然得到

$$(A_1 \oplus B_1)(A_2 \oplus B_2) = (A_1 A_2) \oplus (B_1 B_2) \tag{13}$$

直和空间 $V \oplus W$ 一共有 $(n+m)^2$ 个线性无关矩阵，而子空间 V 和 W 分别只有 n^2, m^2 个线性无关矩阵，这说明并不是所有在直和空间中的矩阵都能表示成直和形式。
利用分块矩阵行列式规则显然得到

$$\det(A \oplus B) = (\det A)(\det B) \tag{14}$$

分块矩阵求迹得到的是各个子块的迹之和

$$\mathrm{Tr}(A \oplus B) = (\mathrm{Tr}\, A) + (\mathrm{Tr}\, B) \tag{15}$$

2 直积

直积这个词在有些书上叫张量积、Kronecker 积

和之前一样，给定两个线性空间 V, W，分别 n, m 维. V 中有一组基底 $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$，$W$ 中也有一组基底 $\left\{\vec{f}_1, \vec{f}_2, \cdots, \vec{f}_m\right\}$. 我们现在定义 nm 个基向量 $\vec{e}_i \otimes \vec{f}_j$，其中 $i = 1, \cdots, n$ 和 $j = 1, \cdots, m$.$\otimes$ 具体运算下面将会讲明，可以看到 $\vec{e}_i \otimes \vec{e}_j$ 有两个指标，这也是为什么直积也叫张量积的原因。

张量积是双线性的，即在 V 中是线性的，在 W 中也是线性的. 如果有多于两个线性空间，则叫多重线性的。这意味着 $\vec{v} \otimes \vec{w} = \left(\sum_i^n v_i \vec{e}_i\right) \otimes \left(\sum_j^m w_j \vec{f}_j\right) = \sum_i^n \sum_j^m v_i w_j \left(\vec{e}_i \otimes \vec{f}_j\right)$. 换句话说由向量基底 $\vec{e}_i \otimes \vec{f}_i$ 生成的线性空间 $V \otimes W$ 中的向量由这组基 $\vec{v}_i \otimes \vec{w}_j$ 线性表出，系数为 $v_i w_j$

一种使它非常明确的方法是把它写出来。例如假设 $n = 2, m = 3$。张量积空间的维数为 $nm = 6$ 新的基底分别是 $\vec{e}_1 \otimes \vec{f}_1, \vec{e}_1 \otimes \vec{f}_2, \vec{e}_1 \otimes \vec{f}_3, \vec{e}_2 \otimes \vec{f}_1, \vec{e}_2 \otimes \vec{f}_2, \vec{e}_2 \otimes \vec{f}_3$. 写出这 6 个基底对应的列向量

$$\begin{gathered}
\vec{e}_1 \otimes \vec{f}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \hline 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_1 \otimes \vec{f}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \hline 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_1 \otimes \vec{f}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \hline 0 \\ 0 \\ 0 \end{pmatrix} \\
\vec{e}_2 \otimes \vec{f}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \hline 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 \otimes \vec{f}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \hline 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e}_2 \otimes \vec{f}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \hline 0 \\ 0 \\ 1 \end{pmatrix}
\end{gathered} \tag{16}$$

这里横线把基向量分割成了两个不同的子集，第一部分由 $\vec{e}_1$ 生成，第二部分由 $\vec{e}_2$ 生成。对于一般的向量 $\vec{v} \in V, \vec{w} \in W$，张量积为

$$\vec{v} \otimes \vec{w} = \begin{pmatrix} v_1 w_1 \\ v_1 w_2 \\ v_1 w_3 \\ \hline v_2 w_1 \\ v_2 w_2 \\ v_2 w_3 \end{pmatrix} \tag{17}$$

在数学上矩阵是一个线性空间到另一个线性空间的线性映射。同样的我们只关注量子力学感兴趣的 $A : V \mapsto V$

$$\vec{v} \longmapsto A\vec{v} \tag{18}$$

或者说

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \longmapsto \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{22} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}_V \tag{19}$$

在张量积空间上，相同的矩阵仍然可以作用在矢量上，使得 $\vec{v} \mapsto A\vec{v}$ 但 $\vec{w} \mapsto \vec{w}$ 保持不变。这个矩阵可以写成 $A \otimes I$，其中 I 是恒等矩阵。在之前的 $n = 2, m = 3$ 例子里，矩阵 A 是 2×2 维矩阵，显然 $A \otimes I$ 是 6×6 维矩阵

$$A \otimes I = \left(\begin{array}{ccc|ccc} a_{11} & 0 & 0 & a_{12} & 0 & 0 \\ 0 & a_{11} & 0 & 0 & a_{12} & 0 \\ 0 & 0 & a_{11} & 0 & 0 & a_{12} \\ \hline a_{21} & 0 & 0 & a_{22} & 0 & 0 \\ 0 & a_{21} & 0 & 0 & a_{22} & 0 \\ 0 & 0 & a_{21} & 0 & 0 & a_{22} \end{array}\right) \tag{20}$$

把这个矩阵作用在 $\vec{v} \otimes \vec{w}$ 上就知道了之所以这样构造的原因

$$\begin{aligned}(A \otimes I)(\vec{v} \otimes \vec{w}) &= \left(\begin{array}{ccc|ccc} a_{11} & 0 & 0 & a_{12} & 0 & 0 \\ 0 & a_{11} & 0 & 0 & a_{12} & 0 \\ 0 & 0 & a_{11} & 0 & 0 & a_{12} \\ \hline a_{21} & 0 & 0 & a_{22} & 0 & 0 \\ 0 & a_{21} & 0 & 0 & a_{22} & 0 \\ 0 & 0 & a_{21} & 0 & 0 & a_{22} \end{array}\right) \left(\begin{array}{c} v_1 w_1 \\ v_1 w_2 \\ v_1 w_3 \\ \hline v_2 w_1 \\ v_2 w_2 \\ v_2 w_3 \end{array}\right) \\ &= \left(\begin{array}{c} (a_{11} v_1 + a_{12} v_2)\, w_1 \\ (a_{11} v_1 + a_{12} v_2)\, w_2 \\ (a_{11} v_1 + a_{12} v_2)\, w_3 \\ \hline (a_{21} v_1 + a_{22} v_2)\, w_1 \\ (a_{21} v_1 + a_{22} v_2)\, w_2 \\ (a_{21} v_1 + a_{22} v_2)\, w_3 \end{array}\right) = (A\vec{v}) \otimes \vec{w}\end{aligned} \tag{21}$$

显然矩阵 A 仅仅只作用在 $\vec{v} \in V$ 上并且保持 $\vec{w} \in W$ 不变。

类似的对于矩阵 $B : W \mapsto W$, $\vec{w} \mapsto B\vec{w}$. 同样的也能用 $I \otimes B$ 作用在 $V \otimes W$ 中的矢量。其中

$$I \otimes B = \left(\begin{array}{ccc|ccc} b_{11} & b_{12} & b_{13} & 0 & 0 & 0 \\ b_{21} & b_{22} & b_{23} & 0 & 0 & 0 \\ b_{31} & b_{32} & b_{33} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & b_{11} & b_{12} & b_{13} \\ 0 & 0 & 0 & b_{21} & b_{22} & b_{23} \\ 0 & 0 & 0 & b_{31} & b_{32} & b_{33} \end{array}\right) \tag{22}$$

作用在 $\vec{v} \otimes \vec{w}$ 为

$$
\begin{aligned}
(I \otimes B)(\vec{v} \otimes \vec{w}) &= \left(\begin{array}{ccc|ccc}
b_{11} & b_{12} & b_{13} & 0 & 0 & 0 \\
b_{21} & b_{22} & b_{23} & 0 & 0 & 0 \\
b_{31} & b_{32} & b_{33} & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & b_{11} & b_{12} & b_{13} \\
0 & 0 & 0 & b_{21} & b_{22} & b_{23} \\
0 & 0 & 0 & b_{31} & b_{32} & b_{33}
\end{array}\right)\left(\begin{array}{c}
v_1 w_1 \\
v_1 w_2 \\
v_1 w_3 \\
\hline
v_2 w_1 \\
v_2 w_2 \\
v_2 w_3
\end{array}\right) \\
&= \left(\begin{array}{c}
v_1\left(b_{11} w_1+b_{12} w_2+b_{13} w_3\right) \\
v_1\left(b_{21} w_1+b_{22} w_2+b_{23} w_3\right) \\
v_1\left(b_{31} w_1+b_{32} w_2+b_{33} w_3\right) \\
\hline
v_2\left(b_{11} w_1+b_{12} w_2+b_{13} w_3\right) \\
v_2\left(b_{21} w_1+b_{22} w_2+b_{23} w_3\right) \\
v_2\left(b_{31} w_1+b_{32} w_2+b_{33} w_3\right)
\end{array}\right) = \vec{v} \otimes (B \vec{w})
\end{aligned} \tag{23}
$$

一般而言

$$
\begin{aligned}
(A \otimes I)(\vec{v} \otimes \vec{w}) &= (A \vec{v}) \otimes \vec{w} \\
(I \otimes B)(\vec{v} \otimes \vec{w}) &= \vec{v} \otimes (B \vec{w})
\end{aligned} \tag{24}
$$

利用上面式子很容易证明下列式子

$$
\begin{aligned}
&\left(A_1 \otimes I\right)\left(A_2 \otimes I\right) = \left(A_1 A_2\right) \otimes I \\
&\left(I \otimes B_1\right)\left(I \otimes B_2\right) = I \otimes \left(B_1 B_2\right) \\
&\left(A_1 \otimes B_1\right)\left(A_2 \otimes B_2\right) = \left(A_1 A_2\right) \otimes \left(B_1 B_2\right) \\
&(A \otimes I)(I \otimes B) = (I \otimes B)(A \otimes I) = (A \otimes B)
\end{aligned} \tag{25}
$$

证明.

$$
(A_1 \otimes I)(A_2 \otimes I)(\vec{v} \otimes \vec{w}) = (A_1 \otimes I)(A_2 \vec{v} \otimes \vec{w}) = (A_1 A_2 \vec{v} \otimes \vec{w}) = ((A_1 A_2) \otimes I)(\vec{v} \otimes \vec{w}) \tag{26}
$$

第二三式类似证明，不再赘述

$$
(A \otimes I)(I \otimes B)(\vec{v} \otimes \vec{w}) = (A \otimes I)(v \otimes B \vec{w}) = (A \vec{v}) \otimes (B \vec{w}) = (A \otimes B)(\vec{v} \otimes \vec{w}) = (I \otimes B)(A \otimes I)(\vec{v} \otimes \vec{w}) \tag{27}
$$

$\square$

最后一式定义了 $A \otimes B$，也可以写出矩阵形式

$$
\begin{aligned}
A \otimes B &= (A \otimes I)(I \otimes B) \\
&= \left(\begin{array}{ccc|ccc}
a_{11} & 0 & 0 & a_{12} & 0 & 0 \\
0 & a_{11} & 0 & 0 & a_{12} & 0 \\
0 & 0 & a_{11} & 0 & 0 & a_{12} \\
\hline
a_{21} & 0 & 0 & a_{22} & 0 & 0 \\
0 & a_{21} & 0 & 0 & a_{22} & 0 \\
0 & 0 & a_{21} & 0 & 0 & a_{22}
\end{array}\right)\left(\begin{array}{ccc|ccc}
b_{11} & b_{12} & b_{13} & 0 & 0 & 0 \\
b_{21} & b_{22} & b_{23} & 0 & 0 & 0 \\
b_{31} & b_{32} & b_{33} & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & b_{11} & b_{12} & b_{13} \\
0 & 0 & 0 & b_{21} & b_{22} & b_{23} \\
0 & 0 & 0 & b_{31} & b_{32} & b_{33}
\end{array}\right) \\
&= \left(\begin{array}{ccc|ccc}
a_{11} b_{11} & a_{11} b_{12} & a_{11} b_{13} & a_{12} b_{11} & a_{12} b_{12} & a_{12} b_{13} \\
a_{11} b_{21} & a_{11} b_{22} & a_{11} b_{23} & a_{12} b_{21} & a_{12} b_{22} & a_{12} b_{23} \\
a_{11} b_{31} & a_{11} b_{32} & a_{11} b_{33} & a_{12} b_{31} & a_{12} b_{32} & a_{12} b_{33} \\
\hline
a_{21} b_{11} & a_{21} b_{12} & a_{21} b_{13} & a_{22} b_{11} & a_{22} b_{12} & a_{22} b_{13} \\
a_{21} b_{21} & a_{21} b_{22} & a_{21} b_{23} & a_{22} b_{21} & a_{22} b_{22} & a_{22} b_{23} \\
a_{21} b_{31} & a_{21} b_{32} & a_{21} b_{33} & a_{22} b_{31} & a_{22} b_{32} & a_{22} b_{33}
\end{array}\right)
\end{aligned} \tag{28}
$$

值得注意，张量积空间中不是所有矩阵都能写成两个子空间矩阵张量积的形式。此外还有两个常用公式

$$\begin{aligned} &\det(A \otimes B) = (\det A)^m (\det B)^n \\ &\operatorname{Tr}(A \otimes B) = (\operatorname{Tr} A)(\operatorname{Tr} B) \end{aligned} \tag{29}$$