集智学园范畴论讲义 v1.7.0809

第二季第 7 课：张量积 张量积的概念遍布数学从具体到抽象的许多方面，从线性代数中的双线性映射，微分几何和物理中的张量积，到模的张量积，模范畴的张量函子，张量的概念会逐渐抽象。范畴论将张量积视为一种泛性质，也视为 Hom 函子的伴随函子。之后的幺半范畴、单子也以张量为基本的例子。这些概念都需要对张量积有实际的认识。

范畴论研究箭头，箭头的并行化组合是基本的问题。从矩阵的箭头方式，讨论了直和方式的并行化。线性映射的并行化在矩阵表示下构成了分块对角矩阵。直和的问题在于，它不是双线性映射，无法纳入到线性空间范畴中考虑，需要张量积来构造性质更好的双线性映射。

接下来用一个常用的例子，复的 Hermite 内积来阐述生活中的旋转张量。这个例子统一了平面上通常的内积、外积的几何意义，并且用双线性的方式将其统一考虑。用 Euler 公式给出了复的版本，用复的双线性的 Hermite 内积统一起来。在基的变换下，观察到旋转不变的结构，并用张量积的方式描述。

最后，在线性空间范畴中，用中介态射的方式，描述了双线性映射的因子化问题，其中的泛性质就是张量积。

直和 态射集 $\mathcal{C}(X,Y)$ 中若定义了加法，则可以将箭头叠加产生等效的效果. 如果是从不同态射集 $\mathcal{C}(X_1,Y_1)$ 和 $\mathcal{C}(X_2,Y_2)$ 中取箭头组 (f_1,f_2)，需要构造某种并行化 $f_1\odot f_2\in\mathcal{C}(X_1\odot X_2,Y_1\odot Y_2)$. 当这样的构造存在，则构成了一种并行化，它是双函子：

$$
\begin{aligned}
\odot:\mathcal{C}\times\mathcal{C}&\to\mathcal{C}\\
(X_1,Y_1)&\mapsto X_1\odot Y_1\\
(X_2,Y_2)&\mapsto X_2\odot Y_2\\
\odot:\mathrm{Hom}((X_1,Y_1),(X_2,Y_2))&\to\mathrm{Hom}(X_1\odot Y_1,X_2\odot Y_2)\\
(f_1,f_2)&\mapsto f_1\odot f_2\\
f_1\odot f_2:X_1\odot Y_1&\to X_2\odot Y_2
\end{aligned}
\tag{0.0.1}
$$

在许多范畴上都可以通过直和来构造这样的并行化：

两个数相互独立，如下图用平行箭头表示. 这里使用了直和记号，n-维列空间可以视为 n 个 $\mathbb{R}$ 的直和 $\mathbb{R}_n=\underbrace{\mathbb{R}\oplus\cdots\oplus\mathbb{R}}_{n\text{ times}}$.

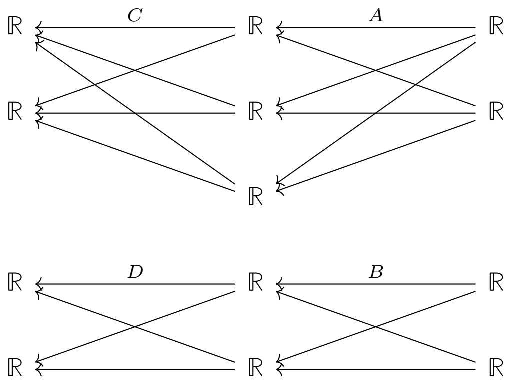
$$
\begin{array}{c}
\mathbb{R}\xleftarrow[a_1]{}\mathbb{R}\\
\\
\mathbb{R}\xleftarrow[a_2]{}\mathbb{R}
\end{array}
\quad=\quad
\mathbb{R}\oplus\mathbb{R}\xleftarrow[a_1\oplus a_2]{}\mathbb{R}\oplus\mathbb{R}
$$

这里 $\bigoplus_{i=1}^n a_i\in\mathrm{End}(\mathbb{R}_n)$ 是 n-维的线性自态射. 前面用实数所代表的线性映射，通过线性的组合构造了矩阵及其所代表的多元线性映射. 类似的，这个构造过程还可以进一步推广到多元线性映射的组合构造，即**直和 (direct sum)**. 进一步如果有两个完全不同的矩阵 $A\in\mathbb{R}_m^n$ 和 $B\in\mathbb{R}_p^q$，二者也可以并行化：

$$
\begin{array}{c}
\mathbb{R}_m\xleftarrow{A}\mathbb{R}_n\\
\\
\mathbb{R}_p\xleftarrow[B]{}\mathbb{R}_q
\end{array}
\quad=\quad
\mathbb{R}_m\oplus\mathbb{R}_p\xleftarrow[A\oplus B]{}\mathbb{R}_n\oplus\mathbb{R}_q
$$

这样可以更好理解分块矩阵的意义. 下图给出了更复杂的流水线:

$$\mathbb{R}_2 \oplus \mathbb{R}_2 \xleftarrow{C\oplus D} \mathbb{R}_3 \oplus \mathbb{R}_2 \xleftarrow{A\oplus B} \mathbb{R}_2 \oplus \mathbb{R}_2$$



分块矩阵的乘法规则相似:

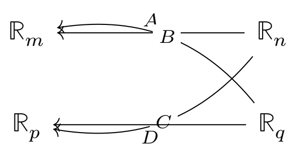
$$\begin{bmatrix} C & \\ & D \end{bmatrix} \circ \begin{bmatrix} A & \\ & B \end{bmatrix} = \begin{bmatrix} C \circ A & \\ & D \circ B \end{bmatrix}$$

这样易于理解直和的复合规律

$$(C \oplus D) \circ (A \oplus B) = (C \circ A) \oplus (D \circ B)$$

对于一般性的非分块对角矩阵的情况

$$\mathbb{R}_m \oplus \mathbb{R}_p \xleftarrow{M} \mathbb{R}_n \oplus \mathbb{R}_q$$



对应的分块矩阵是:

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

直和 $\oplus$ 作为双函子的问题在于，它不是双线性的，因而需要引入新的结构，即张量积. 先从熟悉的线性空间范畴讨论:

线性空间的张量积　讨论实或者复的平面，在许多代数结构上可以认为 $\mathbb{C} \simeq \mathbb{R}^2$. 考虑平面上有向量 $\mathbf{a}$ 和 $\mathbf{b}$，角度分别为 α 和 β，夹角 $\theta = \alpha - \beta$。约定平面实向量在自然基中有坐标:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = a \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = b \begin{bmatrix} \cos\beta \\ \sin\beta \end{bmatrix}$$

这里用黑体强调向量，用正常字体记向量的长度，两向量形成的平行四边形，其夹角为 $\theta = \alpha - \beta$，其最大面积为 ab. 平面内积和外积的几何意义进一步体现为平方和 $(\mathbf{a} \cdot \mathbf{b})^2 + |\mathbf{a} \times \mathbf{b}|^2 = (ab)^2$，着最大面积到内积和外积的正交投影。

$$\begin{bmatrix} \mathbf{a} \cdot \mathbf{b} \\ |\mathbf{a} \times \mathbf{b}| \end{bmatrix} = ab \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = ab \begin{bmatrix} \cos\alpha\cos\beta + \sin\alpha\sin\beta \\ \sin\alpha\cos\beta - \cos\alpha\sin\beta \end{bmatrix} = \begin{bmatrix} a_1b_1 + a_2b_2 \\ -a_1b_2 + a_2b_1 \end{bmatrix} \tag{0.0.2}$$

这里用到了三角函数公式，下面几条性质解释了为何同时考虑内积和外积：

- 两向量张成的最大面积 ab，正交投影分解到内积和外积上，体现出面积是某种旋转不变量；
- 线性规律，内积和外积对于 a 和 b 是双线性的；
- 旋转时，内积和外积都只与夹角 $\theta = \alpha - \beta$ 相关，其线性性来源于三角函数的线性性.

在同构的实/复平面 $\mathbb{C} \simeq \mathbb{R}^2$ 上，可以用复数来表示平面向量，并把以上讨论归纳到复向量之间的 **Hermite 内积 (Hermite inner product)**以便统一研究，根据式**??**有：

$$\begin{aligned}
&\mathbb{C} \times \mathbb{C} \to \mathbb{C} \\
&(\mathbf{a}, \mathbf{b}) \mapsto \langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a} \cdot \overline{\mathbf{b}} \\
&\qquad\quad = ab\mathrm{e}^{i(\alpha-\beta)} = (a\mathrm{e}^{i\alpha})(b\mathrm{e}^{-i\beta}) = ab\mathrm{e}^{i\theta} \\
&\qquad\quad = ab(\cos\alpha + i\sin\alpha) \cdot (\cos\beta - i\sin\beta) \\
&\qquad\quad = ab(\cos\alpha\cos\beta + \sin\alpha\sin\beta) \\
&\qquad\quad + iab(-\cos\alpha\sin\beta + \sin\alpha\cos\beta) \\
&\qquad\quad = (\mathbf{a} \cdot \mathbf{b}) + i(\mathbf{a} \times \mathbf{b})
\end{aligned}$$

其关键之处在于对第二个变元是**共轭 (conjugate)**的，从而在复数乘法中，描述了两复向量的夹角 $\theta = \alpha - \beta$. 式0.0.2中内积和外积的大小由于 2-齐次项构成，即从两向量各取分量对 a_ib_j 形式的项做线性组合。这里的 $\mathbf{a}$ 和 $\mathbf{b}$ 从宏观上代表了双线性映射的两个变元，2-齐次项则从微观上构成了按坐标计算的 2-线性映射的细节. 将 a_ib_j 形式 2-齐次项的所有组合用一个向量表示，构成了接下来讨论的张量：

$$\mathbf{a} \otimes \mathbf{b} = \begin{bmatrix} a_1b_1 \\ a_1b_2 \\ a_2b_1 \\ a_2b_2 \end{bmatrix} = \begin{bmatrix} \cos\alpha\cos\beta \\ \cos\alpha\sin\beta \\ \sin\alpha\cos\beta \\ \sin\alpha\sin\beta \end{bmatrix}$$

内积和外积都只与夹角 $\varphi = \alpha - \beta$ 有关，和绝对的角度 α, β 无关. 现在考虑保持夹角的旋转问题。若 $\mathbb{R}^2$ 有两个正交归一基 $E = [i|j], \tilde{E} = [\tilde{i}|\tilde{j}]$，通过平面旋转变换 $R_\omega \in SO(2)$ 有如下基变换规律：

$$\tilde{E} = R_\omega E = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix} E \tag{0.0.3}$$

那么向量的坐标会发生逆变 $\tilde{\mathbf{a}} = R_\omega^{-1}\mathbf{a}$, $\tilde{\mathbf{b}} = R_\omega^{-1}\mathbf{b}$，展开为：

$$\begin{bmatrix} \tilde{a}_1 \\ \tilde{a}_2 \end{bmatrix} = R_\omega^{-1} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

$$\begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{bmatrix} = R_\theta^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

根据上述逆变关系，考虑 2-齐次向量 $v_{ab} \mapsto \tilde{v}_{ab}$ 是如何随基的变换而变换的. 直接展开有如下关系：

$$\begin{bmatrix} \tilde{a}_1\tilde{b}_1 \\ \tilde{a}_1\tilde{b}_2 \\ \tilde{a}_2\tilde{b}_1 \\ \tilde{a}_2\tilde{b}_2 \end{bmatrix} = \begin{bmatrix} \cos\omega\cos\omega & \cos\omega\sin\omega & \sin\omega\cos\omega & \sin\omega\sin\omega \\ -\cos\omega\sin\omega & \cos\omega\cos\omega & -\sin\omega\sin\omega & \sin\omega\cos\omega \\ -\sin\omega\cos\omega & -\sin\omega\sin\omega & \cos\omega\cos\omega & \cos\omega\sin\omega \\ \sin\omega\sin\omega & -\sin\omega\cos\omega & -\cos\omega\sin\omega & \cos\omega\cos\omega \end{bmatrix} \begin{bmatrix} a_1b_1 \\ a_1b_2 \\ a_2b_1 \\ a_2b_2 \end{bmatrix}$$

中间的这个 4×4 变换矩阵表示了张量积 $R_\omega^{-1}\otimes R_\omega^{-1}$. 综合以上讨论即构成线性映射：

$$\begin{aligned} R_\omega^{-1}\otimes R_\omega^{-1}\ \mathbb{R}^2\otimes\mathbb{R}^2 &\to \mathbb{R}^2\otimes\mathbb{R}^2 \\ \mathbf{a}\otimes\mathbf{b} &\mapsto \tilde{\mathbf{a}}\otimes\tilde{\mathbf{b}} \\ &= (R_\omega^{-1}\otimes R_\omega^{-1})(\mathbf{a}\otimes\mathbf{b}) \end{aligned}$$

映射 $R_\omega^{-1}\otimes R_\omega^{-1}$ 称为平面特殊正交群/旋转群 SO(2) 的 2 阶**逆变张量 (contravariant tensor)**. Hermite 内作为内积是双线性映射：

$$\begin{aligned} H:\mathbb{R}^2\otimes\mathbb{R}^2 &\to \mathbb{C} \\ \mathbf{a}\otimes\mathbf{b} &\mapsto H(\mathbf{a}\otimes\mathbf{b}) \end{aligned}$$

当基变换时，$\mathbf{a}\otimes\mathbf{b}\mapsto\tilde{\mathbf{a}}\otimes\tilde{\mathbf{b}}$ 按照张量 $R_\theta^{-1}\otimes R_\theta^{-1}$ 变换，满足 $H(\tilde{\mathbf{a}}\otimes\tilde{\mathbf{b}}) = H\big((R_\theta^{-1}\otimes R_\theta^{-1})(\mathbf{a}\otimes\mathbf{b})\big) = H(\mathbf{a}\otimes\mathbf{b})$。Hermite 内积、实内积、实外积都只和夹角 $\omega=\alpha-\beta$ 有关，而和两个向量 $\mathbf{a},\mathbf{b}$ 的绝对角度无关. 经过基变换和相应的二阶逆变张量变换 $R_\omega^{-1}\otimes R_\omega^{-1}$ 后，新的张量积 $\tilde{\mathbf{a}}\otimes\tilde{\mathbf{b}}$ 仍然只和夹角 $\theta=\alpha-\beta$ 有关，而和两个向量 $\mathbf{a},\mathbf{b}$ 的绝对角度无关. 实际上这里反映出了 $R_\omega\in SO(2)$ 对两个向量夹角的保持——平面旋转作用于基，使得向量 $\mathbf{a},\mathbf{b}$ 在基下的绝对角度产生了变换，然而平面旋转是正交变换，不改变两向量的夹角，所以它们的相对角度不变.

K-线性空间范畴 $\mathbf{Vct}_K$ 中的张量积，用范畴化的方式定义为：

定义 0.0.1: 线性空间的张量积

在不混淆的情况下，往往记 $\otimes=\otimes_K$. 张量积的定义要求线性空间范畴 $\mathbf{Vct}_K$ 中的下图交换：

$$\begin{tikzcd} X\times Y \arrow[r, "\otimes"] \arrow[dr, "f"'] & X\otimes Y \arrow[d, dashed, "\exists! h"] \\ & Z \end{tikzcd}$$