第 2 章 求和指标、矩阵与张量

刘元彻　　March 9, 2024

在经典电动力学，特别是狭义相对论的学习中，我们经常会遇到基于 Einstein 求和约定进行的“张量”运算——打引号的原因是，有些计算过程中我们遇到的运算根本不是对张量进行的，而只是一些形式上的指标运算。这篇文章目的在于向大家介绍一种自洽的符号体系（convention），用于我们在电动力学的计算中。需要注意的是，这种体系与其相应的、对符号的理解，并不是唯一的，可以参见知乎答主東雲正澍的文章。

这篇文章将按照如下形式组织：第一节中我们将介绍我们的符号规定并给与（物理人可以接受的）基本解释，第二节中我们解释如何将求和符号写成矩阵乘法，第三节中我们将阐述张量、矩阵等一些概念的区别，并解答一些常见问题；附录中将包含一个更加数学化的定义，通过介绍矢量空间及其对偶空间相关的理论，来给出对这套规则更加严谨的理解。

如无特殊声明，本文使用的度规为：$\eta_{\mu\nu} = \mathrm{diag}(-1,+1,+1,+1)$。所有的希腊字母指标 $\mu,\nu,\ldots$ 取值范围为 $0,1,2,3$，所有的拉丁字母指标 $i,j,\ldots$ 取值范围为 $1,2,3$.

2.1 符号规则

首先，我们引入 (1+3)-D　Minkowski 空间的度规：

$$\eta_{\mu\nu} = \begin{cases} -1 & \mu=\nu=0 \\ 1 & \mu=\nu=1,2,3 \\ 0 & \text{otherwise} \end{cases}$$

请注意，这里我们刻意没有将度规写成一个 4×4 的矩阵。这是意在提醒大家，$g_{\mu\nu}$ 作为度规张量的一个分量，是一个数。想要得到一个矩阵，我们提议去掉指标：

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

另外，度规矩阵有一特殊性质：其逆，被定义为两个上标的形式：

$$\eta^{\mu\nu} = \begin{cases} -1 & \mu=\nu=0 \\ 1 & \mu=\nu=1,2,3 \\ 0 & \text{otherwise} \end{cases}$$

也可以写出矩阵形式：

$$\eta^{-1} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

这个定义看起来和度规矩阵本身一模一样，这是 Minkowski 空间中直角坐标系特有的性质。一般的度规性质并不是这样，比如说我们把直角坐标换为球坐标，此时的度规 h 应该写作：

$$h = \mathrm{diag}(-1,1,r^2,r^2\sin^2\theta) \Leftrightarrow \mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2 d\theta^2 + r^2\sin^2\theta \mathrm{d}\varphi^2$$

他的逆就要写为：

$$h^{-1} = (-1,1,\frac{1}{r^2},\frac{1}{r^2\sin^2\theta})$$

现在我们回到 Minkowski 空间。我们现在规定，Minkowski 空间中可以有任意一个带若干指标的数表 $T^{\mu_1\mu_2\cdots\mu_m}_{\nu_1\nu_2\ldots\nu_n}$. 为了保证指标的顺序，我们不接受这种重叠的写法，而是写为（举个例子）：

$$T^{\mu_1\mu_2\ldots\mu_m}{}_{\nu_1\nu_2\ldots\nu_n}, \qquad T_{\mu_1\mu_2\ldots\mu_m}{}^{\nu_1\nu_2\ldots\nu_n}, \qquad T^{\mu_1\mu_2}{}_{\mu_3\ldots\mu_m}{}^{\nu_1\nu_2\nu_3}{}_{\nu_4\ldots\nu_n}\cdots$$

在我们的符号规定中，数表的指标**左右顺序和上下位置都有意义**。我们**不允许**调换一个数表指标的左右顺序，除非对这两个指标进行转置，这需要在矩阵符号上加以体现：

$$F_{\alpha\beta} = [F^T]_{\beta\alpha}, \qquad G_\alpha{}^\beta = [G^T]^\beta{}_\alpha$$

这里调换位置的过程中，指标上下位置是不变的，而矩阵的"名字"上要加一个 T 来表示转置过后的矩阵。

指标的上下改变需要借助度规进行。引入如下的升降指标规则：

$$F^{\mu\nu} = \eta^{\mu\alpha}F_\alpha{}^\nu = \eta^{\mu\alpha}F_{\alpha\beta}\eta^{\beta\nu}$$

在这里我们使用了 Einstein 求和约定，即出现两次的指标将遍历其取值范围，进行求和。

2.2 指标和矩阵的转换

上面的部分介绍了求和约定使用的基本规则。有些同学的疑问是，如何将求和约定转化为一般的矩阵等式？为了正确地写出矩阵等式，我们还要额外引入如下规定。

1. 任何时候，在一个式子等号的同一侧，不允许出现两个同处在上标/下标的相同指标。

   作为例子，以下式子是合法的：

   $$F_{\alpha\beta}F^{\alpha\beta}, \qquad J^\mu A_\mu, \qquad \eta^{\beta\gamma}R^\alpha{}_{\beta\gamma\sigma}$$

   而以下式子不合法：

   $$F_{\alpha\beta}F_{\alpha\beta}, \qquad T_{\mu\nu}J^\mu g_{\mu\alpha}$$

2. 我们仅对相邻的相同指标求和。如果存在可以求和但不相邻的指标，我们可以通过取对两个指标转置或交换被求和量顺序[1]来让他们相邻，例如：

   $$T_{\mu\nu}J^\mu \to J^\mu T_{\mu\mu}, \qquad F^{\mu\nu}F_{\mu\nu} \to [F^T]^{\nu\mu}F_{\mu\nu}$$

   这里最后的 ν 还可以和第一个 ν 求和，但我们无需再改动了，默认最后一个数可以和第一个求和就好。

回忆矩阵乘法写成分量的形式：

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} \to (AB)_{ij} = A_{ik}B_{kj}$$

我们看到，想要改成矩阵乘法，必须要让被求和的指标相邻，这解释了我们的第一条规则。我们用课堂上的一个式子来作为例子：

例题 2.1 将下列等式改为矩阵乘法形式：

$$\Lambda^\mu{}_\rho\eta_{\mu\nu}\Lambda^\nu{}_\sigma\eta^{\sigma\alpha} = \delta^\alpha_\rho$$

解 观察这个式子，我们首先发现有两个 μ 指标不在相邻位置，所以我们需要做一个转置：

$$\implies [\Lambda^T]_\rho{}^\mu\eta_{\mu\nu}\Lambda^\nu{}_\sigma\eta^{\sigma\alpha} = \delta^\alpha_\rho$$

接下来，我们根据度规及其逆矩阵的定义，就写出：

$$\Lambda^T\eta\Lambda\eta^{-1} = \mathbb{I}$$

右乘一个度规矩阵，可以改写为更紧凑的形式：

$$\Lambda^T\eta\Lambda = \eta$$

[1] 请注意，我们这里所有带指标的东西都是数，所以允许交换顺序。如果是量子力学中的算符，那么就只能用转置这一种方法

按照上面的规则，我们也可以把一个矩阵乘法改写成指标形式。当然，我们要注意一些矩阵的指标是有固定定义的，比如度规 $\eta_{\mu\nu}$，其逆 $\eta^{\mu\nu}$，Lorentz 变换 $\Lambda^{\mu}{}_{\nu}$ 等，他们的上下位置不能随意改变。比方说：

例题 2.2 将下列矩阵等式改为带指标的形式：

$$\Lambda^{-1} = \eta^{-1}\Lambda^{T}\eta$$

解 我们知道，Lorentz 逆变换 Λ^{-1} 本身也是一种 Lorentz 变换，因此其指标结构应该为 $[\Lambda^{-1}]^{\mu}{}_{\nu}$. 我们将矩阵等式改写时，首先固定第一个指标，中间用相同指标两两相消：

$$\Longrightarrow\ [\Lambda^{-1}]^{\mu}{}_{\nu} = \eta^{\mu\alpha}[\Lambda^{T}]_{\alpha}{}^{\beta}\eta_{\beta\nu}$$

到这里这式子就改写完了。当然，我们可以设法去掉 Λ^{T} 的转置：

$$\Longrightarrow [\Lambda^{-1}]^{\mu}{}_{\nu} = \eta^{\mu\alpha}\Lambda^{\beta}{}_{\alpha}\eta_{\beta\nu}$$
$$\Longrightarrow [\Lambda^{-1}]^{\mu}{}_{\nu} = \eta_{\beta\nu}\Lambda^{\beta}{}_{\alpha}\eta^{\mu\alpha}$$
$$\Longrightarrow [\Lambda^{-1}]^{\mu}{}_{\nu} = \eta_{\nu\beta}\Lambda^{\beta}{}_{\alpha}\eta^{\alpha\mu}$$

已经凑好了求和的形式，但右边的式子 μ 和 ν 是反的。我们可以考虑做一个转置，改写为矩阵形式：

$$\Longrightarrow\ [\Lambda^{-1}]^{\mu}{}_{\nu} = (\eta\Lambda\eta^{-1})^{T}$$

这式子也是正确的（参见课件）。

2.3 请不要将矩阵和张量混为一谈

这一节我们将向大家澄清矩阵（数表）和张量的区别。一般的线性代数教材其实已经告诉大家，一个线性空间[2]V 应该具有一组基：

$$\{e_1, e_2, \ldots, e_n\} \implies V = \mathrm{span}(e_1, e_2, \ldots, e_n) \qquad \mathrm{dim}V = n$$

经典物理中，我们将所有合法的物理量[3]看作张量。在**没有建立坐标系的情况下**，任何张量在线性变换下**都是不变的**[4]。这和我们的直觉不太相符，因为很容易可以知道我们把坐标架旋转之后，测出来的矢量分量发生了改变。正确理解这种“不变性”的方法是，将一个张量看作一个由分量和基组成的整体：

$$v = v^{i}e_{i}, \qquad F = F^{\mu\nu}e_{\mu}\otimes e_{\nu}$$

在第一个例子中，矢量 v（或者可以称为一个一阶逆变张量）实际上有若干个逆变分量 v^{i} 配合上相应的基 e_i 构成。而在第二个例子中，两个基被 $\otimes$ 拼在一起形成了一个复合基：$e_{\mu}\otimes e_{\mu}$，这个张量 F 被称为一个二阶逆变张量。回忆我们在线性代数中学习到的变换规则：

$$v^{i} \to \Lambda^{i}{}_{j}v^{j}, \qquad e_{i} \to e_{j}[\Lambda^{-1}]^{j}{}_{i} \implies v \to \Lambda^{i}{}_{j}v^{j}e_{k}[\Lambda^{-1}]^{k}{}_{i} = \delta^{k}_{i}v^{i}e_{k} = v$$

上面的 Λ 是线性空间 V 到自身的线性变换（不一定是 Minkowski 空间），从这里我们给出两个规则：

1. 逆变矢量的分量就像我们在线性代数里学到的矢量分量一样，当坐标基矢作某个变换时，他们将会作相反的变换，这使得矢量整体保持不变
2. 对更高阶的逆变张量，可以用张量积 $\otimes$ 将一些基并在一起。$\otimes$ 具有结合性，但**不具有交换性**。可以逐个写出变换来验证高阶的逆变张量保持不变。

为什么上面的例子中没有出现带有下标的分量呢？其实是因为他们并不是我们通常说的矢量空间中的元素，而是生活在对偶空间 V^{*} 中的元素。对偶空间[5]也有一组基：

$$\{e^{*1}, e^{*2}, \ldots, e^{*n}\} \implies V^{*} = \mathrm{span}(e^{*1}, e^{*2}, \ldots, e^{*n}) \qquad \mathrm{dim}V^{*} = \mathrm{dim}V = n$$

[2]我们只考虑有限维线性空间

[3]请注意，是物理量的整体。一个物理量比如动量 p 在某个坐标系里的分量并不是我们所说的“物理量”

[4]这在线性代数中常被称为线性变换的被动观点

[5]有限维的对偶空间会比较简单，无限维则需要考虑各种问题，参见泛函分析的内容

所谓对偶空间，实际上是用**将矢量空间映射到数域上的线性变换**构成一个线性空间：

$$w^* \in V^* : V \to \mathbb{K} \implies v \to \omega^*(v)$$

其中我们定义对偶基具有如下性质：

$$e^{*i}(e_j) = \delta^i_j$$

这样我们就拥有了任意一个线性空间中的 Kronecker 符号，特别提醒这个 Kronecker 符号永远具有一上一下的指标，这和空间中的度规是无关的。

对偶矢量也可以像矢量一样构造张量：

$$p = p_i e^{*i}, \qquad G = G_{\mu\nu} e^{*\mu} \otimes e^{*\nu}$$

其中带有下标的系数称为协变分量，他和坐标基矢 e_i 拥有一样的变换规则：

$$p_i \to p_j [\Lambda^{-1}]^j{}_i, \qquad e^{*i} \to \Lambda^i{}_j e^{*j} \implies p \to p$$

我们在实际使用张量时，常常省略他们的基，而只写出分量。既然已经指定基矢，我们得到的分量式就可以和矩阵相互转化，参见前两节。

讲述矢量和对偶矢量的原因是方便澄清数表和张量的区别。例如：

$$\partial_\mu F^{\mu\nu} = -\mu_0 J^\nu$$

仅仅写出这个等式是危险的。因为在我们的符号规则中，我们并不在乎被求和的指标究竟属于数表还是张量。因此，我们写出来的表达式仅仅在数学形式上正确，而不一定具有物理意义。这回答了有些同学提出的问题：Lorentz 变换矩阵是否可以用度规升降指标？回答是可以，但这么做没有物理意义。因为对于张量，他的分量具有**协变性和逆变性**，即一定与坐标基矢作相同变换或相反变换；但对于数表则**没有这种变换规律**。

到这里，我们已经解释了一个带有指标的量究竟应该如何解读。为了方便，我将顺序梳理一下：

1. 首先，先确定这个量是张量还是数表（这很重要！）；
2. 对于数表，他的指标上下位置并没有特殊的意义，仅仅是为了保持形式上的指标平衡；
3. 对于张量，上标表示逆变（即该分量与坐标基矢变换规律相反），下标表示协变（即该分量与坐标基矢变换规律相同）；
4. 无论是张量还是数表，我们都可以通过转置来交换相邻两个指标的左右位置。另外，在带有指标的分量表达式中，被求和的都是数，所以可以交换；
5. 以上步骤完成后，就可以放心地在分量表达式和矩阵表达式中间相互转化。

参考资料

关于求和记号的部分，主要参考杨焕雄老师的课程幻灯片。学界常用的求和规则也有其他形式，感兴趣的同学可以搜索：知乎答主-東雲正澍和 Misaki Notation；Dotted-Undotted 规则等。

关于线性空间、对偶空间和张量的描述，我们采取的记号被称为具体指标，与之相对的被称为抽象指标。这些内容的论述可以参考梁灿彬老师的《微分几何入门与广义相对论（上册）》。