

0.0.1 数域

读者应该已经熟悉自然数、整数、有理数、实数、复数等数集，以及定义在它们之上的四则运算。现在我们特别关心这些数的集合是否具有对四则运算的封闭性，即集合的元素经过四则运算之后的结果仍属于原集合。具体地，设 \mathbb{F} 是数的集合，且在 \mathbb{F} 上的四则运算满足以下要求：

1. 加法交换律： $x + y = y + x, \forall x, y \in \mathbb{F}$
2. 加法结合律： $x + (y + z) = (x + y) + z, \forall x, y, z \in \mathbb{F}$
3. 零： \mathbb{F} 中存在一个零元素 0 ，使得 $x + 0 = x, \forall x \in \mathbb{F}$
4. 对相反数封闭：对任一 $x \in \mathbb{F}$ ，总存在唯一 \mathbb{F} 的元素，记作 $-x$ ，满足 $x + (-x) = 0$
5. 乘法交换律： $xy = yx, \forall x, y \in \mathbb{F}$
6. 乘法结合律： $x(yz) = (xy)z, \forall x, y, z \in \mathbb{F}$
7. 单位元： \mathbb{F} 中存在一个非零元素 1 ，满足 $x1 = x, \forall x \in \mathbb{F}$
8. 对倒数封闭：对 \mathbb{F} 中的任一非零元素 x ，总存在唯一元素，记作 x^{-1} 或 $1/x$ ，满足 $xx^{-1} = 1$
9. 乘法对加法的分配律： $x(y + z) = xy + xz, \forall x, y, z \in \mathbb{F}$

则称集合 \mathbb{F} 是一个域 (*field*)。易验复数 \mathbb{C} 、实数 \mathbb{R} 和有理数 \mathbb{Q} 是域。整数 \mathbb{Z} 、自然数 \mathbb{N} 不是域。如果一个域 \mathbb{F} 的子集 \mathbb{F}' 也是域，则称 \mathbb{F}' 是 \mathbb{F} 的子域 (*subfield*)。例如， \mathbb{C} 是 \mathbb{C} 的子域； \mathbb{R} 是 \mathbb{C} 的子域；所有形如 $x + y\sqrt{2}, x, y \in \mathbb{C}$ 的复数集合是 \mathbb{C} 的子域。

一般地，域的定义不具体规定加法和乘法的具体运算规则。只要一个集合上所谓的加法和所谓的乘法分别满足上列 9 个要求，它就是一个域。在这种一般性下若不作更多规定，将会允许一些奇怪的域，本讲义不深入探讨。在接下来的内容里，每当提到“数域 \mathbb{F} ”，都具体指复数域 \mathbb{C} 的任一子域。

0.0.2 向量空间的定义和基本性质

定义 0.1 (向量空间). 给定一个数域 \mathbb{F} 和一个非空集合 \mathcal{V} ，如果它们满足：

1. \mathcal{V} 的元素的二元运算 (*binary operation*) 规则，称为加法，记为 $\mathbf{a} + \mathbf{b}$ ， $\mathbf{a}, \mathbf{b} \in \mathcal{V}$ 。该运算满足以下规定： $\forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{V}$ ：
 - (a) 封闭性 (*closure*): $\mathbf{a} + \mathbf{b} \in \mathcal{V}$
 - (b) 交换律 (*commutative*)^{*}: $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
 - (c) 结合律 (*associative*)[†]: $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$
 - (d) 单位元 (*identity element*): $\exists \mathbf{0} \in \mathcal{V} : \mathbf{a} + \mathbf{0} = \mathbf{a}$

^{*}对应的英文单词是形容词。

[†]对应的英文单词是形容词。

(e) 逆元 (*inverse element*): $\forall \mathbf{a} \in \mathcal{V}, \exists -\mathbf{a} : \mathbf{a} + (-\mathbf{a}) = \mathbf{0}$

2. 数域 \mathbb{F} 与 \mathcal{V} 的元素间的二元运算规则, 称为标量乘法, 记为 $\alpha \mathbf{a}$, $\alpha \in \mathbb{F}, \mathbf{a} \in \mathcal{V}$ 。该运算满足以下规定: $\forall \alpha, \beta \in \mathbb{F}, \mathbf{a}, \mathbf{b} \in \mathcal{V}$:

(a) $\alpha \mathbf{a} \in \mathcal{V}$

(b) $\alpha(\beta \mathbf{a}) = (\alpha\beta) \mathbf{a}$

(c) $1\mathbf{a} = \mathbf{a}$, 其中 1 是数域 \mathbb{F} 的单位元

(d) $\alpha(\mathbf{a} + \mathbf{b}) = \alpha \mathbf{a} + \alpha \mathbf{b}$

(e) $(\alpha + \beta) \mathbf{a} = \alpha \mathbf{a} + \beta \mathbf{a}$

则称 \mathcal{V} 是数域 \mathbb{F} 上的向量空间 (*vector space*); \mathcal{V} 的元素 $\mathbf{a}, \mathbf{b}, \dots$ 是向量 (*vector*); \mathbb{F} 中的数是 \mathcal{V} 的标量 (*scalar*)。

读者可尝试利用定义证明 $0\mathbf{a} = \mathbf{0}, \forall \mathbf{a} \in \mathcal{V}$ 。由于这条可由定义证明, 故尽管它看上去也很基本, 但并不写在定义中。

在定义 0.1 中, 数域 \mathbb{F} 是关于标量乘的运算规定 (2-a) 中的标量所属的数域。任何非空集合, 只要配备了定义 0.1 所规定的运算法则, 就是一个向量空间, 其元素就是向量。因此, 向量是抽象的一般概念。大一的《线性代数与解析几何》课本中的“行/列向量”和“矩阵”, 都只是特例。例 0.1 给出了更多向量空间的例子。这种抽象的定义方式, 使得只要是在抽象的层面上论证得到的性质和定理, 就必然适用于所有具体的例子当中, 使我们学习和发展数学事半功倍。

定义 0.1 中还有若干需要注意的细节。比如, (1-d) 和 (1-e) 都只规定了存在性, 而没有规定唯一性。其实单位元是唯一的; 相对于一个向量的逆元也是唯一的。这两个命题可以由定义 0.1 证明出来, 故无需添加到定义中作为规定。请回顾《线性代数与解析几何》第七章“2. 线性空间的性质”下的所有内容。

例 0.1. 以下是一些向量空间的例子。

1. \mathbb{R}^n 是所有有序实数 n 元组 $(a_i) = (\alpha_1, \dots, \alpha_n), \alpha_i \in \mathbb{R}, i = 1, \dots, n$ 的集合。若 $\forall (a_i), (b_i) \in \mathbb{R}^n$:

(a) $(a_i) + (b_i) = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)$;

(b) $\alpha(a_i) = (\alpha\alpha_1, \dots, \alpha\alpha_n), \forall \alpha \in \mathbb{R}$;

(c) $(0) = (0, \dots, 0)$;

(d) $-(a_i) = (-\alpha_1, \dots, -\alpha_n)$;

则 \mathbb{R}^n 连同上述的运算规定形成数域 \mathbb{R} 上的一个向量空间, 又称为实坐标空间。符号 \mathbb{R} 或 \mathbb{C} 既可表示数域, 又可表示一个一维实或复坐标空间。可以验证, 一维实坐标空间 \mathbb{R} 不是数域 \mathbb{C} 上的向量空间。

2. 数域 \mathbb{F} 上的所有 $m \times n$ 矩阵的集合 $\mathbb{F}^{m \times n}$ (连同矩阵的加法和矩阵的标量乘法规定*) 是一个向量空间。其零向量是 $m \times n$ 全零矩阵^[1]§7.1 例 1.3。
3. 验证: 在开区间 (a, b) 上的所有实值一元连续函数的集合是实数域上的向量空间^[1]§7.1 例 1.1。
4. 验证: 记 $\mathcal{C}^n(a, b)$ 为开区间 (a, b) 上的所有 n 阶连续可导实值一元函数的集合, 它是实数域上的向量空间。

接下来直至定义0.6, 我们将逐步发现, 定义0.1中的性质和运算法则将进一步导致向量空间有一定的维数。

由于“封闭性”的要求, 一个向量空间内的任一向量总能被这一向量空间中的其他向量按所规定的运算法则表达出来。具体地, 若 \mathcal{V} 是数域 \mathbb{F} 上的向量空间, $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathcal{V}$, $\alpha_1, \dots, \alpha_n \in \mathbb{F}$, 则依据向量空间的封闭性, $\sum_{i=1}^n \alpha_i \mathbf{a}_i$ 也属于 \mathcal{V} , 即这一求和的结果也是一个向量, 该向量可以用前面这个求和表达式来表达。由此引出线性组合、线性表出和线性无关的概念。

定义 0.2 (线性组合、线性表出、线性无关). ^[1]§7.2 定义 2.2、2.3 若 \mathcal{V} 是 \mathbb{F} 上的向量空间, $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathcal{V}$, $\alpha_1, \dots, \alpha_n \in \mathbb{F}$, 则 $\sum_{i=1}^n \alpha_i \mathbf{a}_i$ 称为这 n 个向量 $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ 的线性组合 (linear combination)。令 $\mathbf{b} = \sum_{i=1}^n \alpha_i \mathbf{a}_i$, 则称 \mathbf{b} 被 $\{\mathbf{a}_i\}$ 线性表出[†]。若 $\sum_{i=1}^n \alpha_i \mathbf{a}_i = \mathbf{0}$ 当且仅当 $\alpha_i = 0 \forall i$, 则称向量 $\{\mathbf{a}_i\}$ 线性无关 (linear independent)。反之, 若存在某不全为零的一组 $\{\alpha_i\}$, $\alpha_i \in \mathbb{F}, i = 1, \dots, n$ 使得 $\sum_{i=1}^n \alpha_i \mathbf{a}_i = \mathbf{0}$ 则称向量 $\{\mathbf{a}_i\}$ 线性相关 (linear dependent)。

由定义0.2可证得如下结论^[1]§7.2 定义 2.3 下的“常用结果”:

1. 任何真包含一组线性无关向量的向量集合是线性相关的。
2. 任何线性无关向量组的子集也是线性无关向量组。
3. 任何含有 $\mathbf{0}$ 向量的向量组线性相关, 因为总有 $1 \neq 0$ 使得 $1\mathbf{0} = \mathbf{0}$ 。
4. 一个向量组 S 是线性无关向量组, 当且仅当 S 的所有子集都是线性无关向量组。

向量空间定义中所要求的封闭性保证了向量可类似于我们习惯的数字那样被用作数学表达和运算。因此, 封闭性是一个很重要的性质。那么, 一个向量空间之内, 会不会有一部分子集本身就满足了封闭性呢 (就好像复数与实数之间的关系)? 我们通过考察 \mathbb{C}^n 和 \mathbb{R}^n 可以举出很多正面的例子。一般地, 如果一个向量空间的子集本身也满足封闭性, 那么它自己也是一个向量空间 (即满足定义0.1)。

定义 0.3 (子空间). ^[1]§7.1 定义 1.2 及其后例题] 令 \mathcal{V} 是数域 \mathbb{F} 上的一个向量空间, 如果 \mathcal{V} 的子集 $\mathcal{W} \subseteq \mathcal{V}$ 也是一个向量空间 (并与 \mathcal{V} 采用相同的加法和标量乘定义), 则称 \mathcal{W} 是 \mathcal{V} 的一个子空间 (subspace)。

*见^[1]§2.1 矩阵与矩阵的运算

[†]集合 $\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \subset \mathcal{V}$ 可写为 $\{\mathbf{a}_i\}_{i=1}^n$ 或 $\{\mathbf{a}_i\}$, $\{\mathbf{a}_i\}$ 的一个有序序列则记为 (\mathbf{a}_i) 。

按照这一定义，易证 \mathcal{W} 的任意两个向量 $\mathbf{a}, \mathbf{b} \in \mathcal{W}$ 和任一标量 $\alpha \in \mathbb{F}$ 构成的线性组合 $\alpha\mathbf{a} + \mathbf{b}$ 也在 \mathcal{W} 内^[1]§7.1 定理 1.1。

就算向量空间 \mathcal{V} 的某子集 S 因不满足封闭性而成为不了 \mathcal{V} 的子空间， S 内的向量的所有线性组合表出的向量可以形成一个比 S 更大的集合，且满足封闭性，因而我们可以说由 S 生成了一个 \mathcal{V} 的子空间，定义如下。

定义 0.4 (线性生成空间). 若 S 是向量空间 \mathcal{V} 的非空子集，即 $S \subseteq \mathcal{V}, S \neq \emptyset$ ，那么 S 内的向量的所有线性组合的集合 \mathcal{W}_S 也是一个向量空间，称为由 S 线性生成的子空间 (*the subspace spanned by S*)。

换句话说， \mathcal{W}_S 中的向量都能由 S 的向量线性表出。一个直接的结论就是 $S \subseteq \mathcal{W}_S$ ，因为一个向量总能被它自己线性表出。

例 0.2.

- 易验证，三维实坐标空间 \mathbb{R}^3 的子集 $P = \{(x_1, x_2, x_3) | x_1 = 0, x_2, x_3 \in \mathbb{R}\}$ 具有封闭性，因此它是 \mathbb{R}^3 的子空间。
- 设 Q 是 \mathbb{R}^3 的子集 $\{(2, 1, 3), (1, 0, 1)\}$ 线性生成的子空间，常记为

$$Q = \text{span} \{(2, 1, 3), (1, 0, 1)\}$$

则可验 $\{(7, 2, 9)\} \in Q$ 。

- 如果用有序实数三元组表示从原点 $(0, 0, 0)$ 出发的矢量，则上例中的 $(2, 1, 3), (1, 0, 1)$ 的两个矢量不共线（易验它们线性无关），子空间 Q 是由这两个矢量所确定的平面。

从上面的例子看到，一个向量空间与其子空间之间的关系，暗示了某种维度的概念。我们首先可以明确“一个向量空间维数”的一般意义，但需要先引入“基”的概念。

定义 0.5 (向量空间的基). ^[1]§7.2 定义 2.4 如果向量空间 \mathcal{V} 是其子空间 B 的线性生成空间（即 $\mathcal{V} = \text{span}B$ ，且 B 内的所有向量线性无关，则称 B 是 \mathcal{V} 的一组基 (*basis*)。如果 B 含有有限个向量，则称 \mathcal{V} 是有限维向量空间。

上一定义仅引入了“有限维”的概念，但没有具体涉及到“维数是几”的问题。因为按目前已有的定义， \mathcal{V} 内可能可以找出不止一组满足定义的基，而这些基是否必然都具有相同个数的向量？只有当这个问题的答案是肯定的，我们才能通过把它们个数统一定义为 \mathcal{V} 的维数，来使得向量空间具有确定的维数。下面的定理解决了这个问题。

定理 0.1. 有限维向量空间的每组基具有相同个数的线性无关向量。

证明. 此略^[1]“(3) 的证明”, p. 171^[2]§2.3, Theorem 4, p. 44。

□

有了这一定理, 我们就可以直接把有限维向量空间的维数定义为它的任一组基的向量个数——

定义 0.6 (有限维向量空间的维数). 设 \mathcal{V} 是数域 \mathbb{F} 上的有限维向量空间, 它的维数 (*dimension*), 记为 $\dim \mathcal{V}$, 是它的任一组基的向量个数。规定: 零向量空间 (只有一个零向量组成的向量空间) 是 0 维。

由定理 0.1 还可直接得到如下推论, 它们的证明与定理 0.1 的证明过程很类似, 故从略。

推论 0.1.1. 设 \mathcal{V} 是一个有限维向量空间, 其维数 $\dim \mathcal{V} = n$, 则

1. \mathcal{V} 的任一含有多于 n 个向量的子集都是线性相关向量组。
2. \mathcal{V} 的任一向量个数少于 n 的子集都不能线性生成整个 \mathcal{V} (即这样的子集的线性生成空间总是 \mathcal{V} 的真子集)。
3. \mathcal{V} 的任一子空间 \mathcal{W} 的维数不大于 \mathcal{V} 的维数, 即 $\dim \mathcal{W} \leq \dim \mathcal{V}$; 当且仅当 $\mathcal{W} = \mathcal{V}$ 时取等号。

这一推论的第 3 条的一个例子就是之前的例 0.2。

例 0.3. 验证以下命题——

- 如果把一维实坐标空间 \mathbb{R} 看作是实数域 \mathbb{R} 上的向量空间, 则 $\{1\}$ 是其一组基, 故一维实坐标空间 \mathbb{R} 是实数域 \mathbb{R} 上的一维向量空间。
- 如果把一维复坐标空间 \mathbb{C} 看作是实数域 \mathbb{R} 上的向量空间, 则 $\{1, i\}$ 是其一组基, 故一维复坐标空间 \mathbb{C} 是实数域 \mathbb{R} 上的二维向量空间。
- 如果把一维复坐标空间 \mathbb{C} 看作是复数域 \mathbb{C} 上的向量空间, 则 $\{1\}$ 是其一组基, 故一维复坐标空间 \mathbb{C} 是复数域 \mathbb{C} 上的一维向量空间。

0.0.3 向量的坐标

到此为止, 我们未具体地阐明“向量”是什么, 也未具体地规定加法和标量乘法如何进行。“这种抽象性使我们可以把不同的数学对象统一到线性空间这一概念之下。”^[1]p. 167 不过, 通过引入“坐标”的概念, 我们又使得任一抽象向量都能用一组有序数组来唯一地表示, 从而使抽象的向量之间的运算得以由具体的有序数组的运算来代替 (就像我们以往在《线性代数》课中所熟悉的那样)。

定义 0.7 (向量在给定有序基下的坐标). 若 \mathcal{V} 是一个 n 维向量空间, \mathcal{V} 内的一组线性无关的有序向量序列 (\mathbf{a}_i) 线性生成整个 \mathcal{V} , 则称这组有序向量序列 B 为 \mathcal{V} 的一组有序基 (ordered basis)。由定义 0.5, 任一向量 $\mathbf{x} \in \mathcal{V}$ 均可表达为 $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{a}_i$, 进而, 任一 $\mathbf{x} \in \mathcal{V}$ 在给定有序基下都唯一对应 \mathbb{F}^n 中的一个有序 n 元数组 $(x_i)^*$, 我们称这一有序 n 元数组 (x_i) 为向量 \mathbf{x} 在有序基 B 下的坐标 (coordinate)。

基的原始定义 (0.5) 仅要求基是一个集合, 没有有序性的规定。故我们可以书写 “ $B = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ 是某向量空间的一组基”。但由于集合的元素是无序的, 用一组基向量表出任意一个 n 维向量时所使用的 n 个标量, 若只形成一个集合, 那么这一标量集合的元素变换不同的顺序去与基向量组合, 将线性表出不同的向量。因此我们必须给基向量规定顺序, 形成有序基, 才能说出 “一个向量的第 i 个坐标”。换句话说, 向量坐标的定义需要基于有序基, 而非仅是基向量的一个集合。所以定义 0.7 才特别增加了 “有序” 的要求。在本讲义中, 我们都用同一个符号来表示作为集合而言的一组基 (例如 “ $B = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ 是某向量空间的一组基”) 和一组有序基 (例如 “ $B = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ 是某向量空间的一组有序基”)。

在定义 0.7 中提到的向量 \mathbf{x} 与其在有序基 B 下的坐标 (x_1, \dots, x_n) 之间的唯一对应性, 可以很方便地证明。沿用定义 0.7 中的记法, 若另有某 $(y_1, \dots, y_n) \in \mathbb{F}^n$ 满足 $\mathbf{x} = \sum_{i=1}^n y_i \mathbf{a}_i$, 其中至少有一 $y_k \notin \{x_1, \dots, x_n\}$, 则 $\mathbf{x} - \mathbf{x} = \sum_{i=1}^n (x_i - y_i) \mathbf{a}_i = \mathbf{0}$ 且 $(x_1 - y_1, \dots, x_n - y_n)$ 中有一个数 $x_k - y_k \neq 0$, 与 $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ 线性无关矛盾。故 (x_1, \dots, x_n) 对 \mathbf{x} 是唯一的。

有了坐标的定义, 在给定基下, n 维向量空间 \mathcal{V} 中的每一个向量就都与 \mathbb{F}^n 中的一个有序 n 元数组形成了一一对应的关系。由 \mathcal{V} 和 \mathbb{F} 的封闭性, 没有一个向量不对应一个数组, 反之亦然。易验, 通过向量的加法和标量乘法所得到的新向量所对应的坐标, 就是原向量的坐标按照 \mathbb{F}^n 上的加法和标量乘法运算的结果[†]。但是要注意, n 维实坐标空间 \mathbb{R}^n 中的一个向量 $\mathbf{x} \in \mathbb{R}^n$ 本身就是一个有序实数 n 元组 $\mathbf{x} = (x_1, \dots, x_n), x_i \in \mathbb{R}, i = 1, \dots, n$ 。在选定 \mathbb{F}^n 某组有序基 $B = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ 下 (这些基向量本身也都是有序实数 n 元组), 向量 \mathbf{x} 的坐标可能又是另一个不同的有序实数 n 元组 (χ_1, \dots, χ_n) 。正确的表示是 $\mathbf{x} = \sum_{i=1}^n \chi_i \mathbf{e}_i$, 或称向量 \mathbf{x} 在有序基 B 下的坐标是 (χ_1, \dots, χ_n) , 但不能写成 $\mathbf{x} = (\chi_1, \dots, \chi_n)$ [‡]。

例 0.4. 设 $B = \{\alpha_1, \alpha_2, \alpha_3\}$ 是 \mathbb{R}^3 的一组有序基, 其中

$$\alpha_1 = (1, 0, -1), \quad \alpha_2 = (1, 1, 1), \quad \alpha_3 = (1, 0, 0)$$

请问向量 (a, b, c) 在 B 下的坐标表达式?

*这里的唯一性可参考[1]§7.2 定义 2.3 下 “(3) 的证明”。

† “利用基和坐标可把线性空间的运算变得更具体。”[1]p.173。

‡这里的概念区分可参见[1]§7.1 例题 2.1。

在本讲义中，无论是 n 维向量 \mathbf{x} 本身还是其在某基的下的坐标，都一律写成 $n \times 1$ 矩阵（列向量）；为方便，在文字段落中表示为 $1 \times n$ 矩阵（行向量）的转置，即 $\mathbf{x} = (x_1, \dots, x_n)^\top$ 。这里的转置可直接按以前在《线性代数》课中学过的意义来理解。但本讲义会对向量转置的概念进行正式的定义。

0.0.4 向量的内积与范

在本节中，我们在向量空间的基础上再增加一些性质和运算法则，使得“正交”、“单位向量”的概念得以引入，同时也介绍其他相关的概念。

定义 0.8 (内积). 数域 \mathbb{F} 上的向量空间 \mathcal{V} 中两向量 $\mathbf{a}, \mathbf{b} \in \mathcal{V}$ 的内积 (*inner product*) 记为 $(\mathbf{a}|\mathbf{b})$ ($\cdot|\cdot$) $\mathcal{V} \times \mathcal{V}$ 到 \mathbb{F} 的映射，满足以下规定*： $\forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{V}, \alpha \in \mathbb{F}$

1. $(\mathbf{a}|\mathbf{b}) = \overline{(\mathbf{b}|\mathbf{a})}$
2. $(\alpha \mathbf{a}|\mathbf{b}) = \alpha (\mathbf{a}|\mathbf{b})$
3. $(\mathbf{a} + \mathbf{b}|\mathbf{c}) = (\mathbf{a}|\mathbf{c}) + (\mathbf{b}|\mathbf{c})$
4. $(\mathbf{a}|\mathbf{a}) \in \mathbb{R}$ 且 $(\mathbf{a}|\mathbf{a}) \geq 0$ ，当且仅当 $\mathbf{a} = \mathbf{0}$ 时取等号。

带有一种内积运算规定的向量空间叫做内积空间 (*inner product space*)。

由内积运算要求的第 3 条易知 $\forall \mathbf{a} \in \mathcal{V}, (\mathbf{a}|\mathbf{0}) = 0$ ，因为 $\forall \mathbf{a}, \mathbf{b} \in \mathcal{V}$ 均有 $(\mathbf{a}|\mathbf{b}) = (\mathbf{a}|\mathbf{0} + \mathbf{b}) = (\mathbf{a}|\mathbf{0}) + (\mathbf{a}|\mathbf{b}) \Rightarrow (\mathbf{a}|\mathbf{0}) = 0$ 。由内积定义还可以推出： $\forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{V}, \alpha \in \mathbb{F}$

$$(\mathbf{a}|\alpha \mathbf{b} + \mathbf{c}) = \alpha (\mathbf{a}|\mathbf{b}) + (\mathbf{a}|\mathbf{c})$$

即从内积的后一个向量提出标量到内积之外时，这一标量要取复数共轭。

内积的定义中设置复数共轭是必要的。否则将面临如下的矛盾：由 $(\mathbf{a}|\mathbf{a}) > 0 \forall \mathbf{a} \neq \mathbf{0}$ ，竟然有 $(i\mathbf{a}|i\mathbf{a}) = -1 (\mathbf{a}|\mathbf{a}) > 0$ 。但是，到底是规定从内积的前一个向量提出的标量要取复数共轭（很多物理书习惯），还是从内积的后一个向量提出的标量要取复数共轭（本讲义的定义方式，也是很多数学书的习惯）？这个惯例的不同，有时会造成重要的差别。我们可以把两种惯例的内积定义用不同的括号来区分，使得这两种惯例的内积的关系是：

$$(\mathbf{x}|\mathbf{y}) \equiv \langle \mathbf{y}|\mathbf{x} \rangle$$

上面的三角括号定义与量子力学中的狄拉克 bra-ket 标记的规定是相同的。由于本讲义的流变学部分不涉及复数域上的向量空间，因此不再详述，可参考[3]§1.3.1。

例 0.5.

*上划线表示复数共轭

1. 在 \mathbb{F}^n 上可定义这样的内积：对于 $\mathbf{a} = (\alpha_1, \dots, \alpha_n)^\top, \mathbf{b} = (\beta_1, \dots, \beta_n)^\top \in \mathbb{F}$, $(\mathbf{a}|\mathbf{b}) \equiv \sum_j \alpha_j \overline{\beta_j}$, 称为 \mathbb{F}^n 上的标准内积 (standard inner product)。 \mathbb{R}^n 上的标准内积又可记为点乘 (dot product) $\mathbf{a} \cdot \mathbf{b}$ 。
2. 记 $\mathbb{C}(a, b)$ 为所有定义在实开区间 (a, b) 上的复数值一元函数的集合, 若通过 (a, b) 上的恒正函数 $w(x)$ 定义内积运算 $(f|g) = \int_a^b w(x) f(x) \overline{g(x)} dx, \forall f, g \in \mathbb{C}(a, b)$, 验证 $\mathbb{C}(a, b)$ 是一个内积空间。

由内积的一般定义可直接证明柯西-施瓦茨不等式 (Cauchy-Schwarz inequality), 故该不等式在任一内积空间上都成立。

定理 0.2. 设 \mathcal{V} 是数域 \mathbb{F} 上的一个内积空间, 则有 $\forall \mathbf{a}, \mathbf{b} \in \mathcal{V}$,

1. 柯西-施瓦茨不等式: $|(\mathbf{a}|\mathbf{b})|^2 \leq (\mathbf{a}|\mathbf{a})(\mathbf{b}|\mathbf{b})$
2. $(\mathbf{a}|\mathbf{b}) = \operatorname{Re}(\mathbf{a}|\mathbf{b}) + i \operatorname{Re}(\mathbf{a}|i\mathbf{b})$

证明. 当 $\mathbf{a} = \mathbf{0}$ 时, 不等式取等号成立。当 $\mathbf{a} \neq \mathbf{0}$ 时, 令

$$\mathbf{c} = \mathbf{b} - \frac{(\mathbf{b}|\mathbf{a})}{(\mathbf{a}|\mathbf{a})} \mathbf{a}$$

则可验证 $(\mathbf{c}|\mathbf{a}) = 0$, 且

$$\begin{aligned} 0 \leq (\mathbf{c}|\mathbf{c}) &= \left(\mathbf{b} - \frac{(\mathbf{b}|\mathbf{a})}{(\mathbf{a}|\mathbf{a})} \mathbf{a} \middle| \mathbf{b} - \frac{(\mathbf{b}|\mathbf{a})}{(\mathbf{a}|\mathbf{a})} \mathbf{a} \right) \\ &= (\mathbf{b}|\mathbf{b}) - \frac{|(\mathbf{b}|\mathbf{a})|^2}{(\mathbf{a}|\mathbf{a})} \\ &\Leftrightarrow |(\mathbf{a}|\mathbf{b})|^2 \leq (\mathbf{a}|\mathbf{a})(\mathbf{b}|\mathbf{b}) \end{aligned}$$

由 $(\mathbf{a}|\mathbf{b}) = \operatorname{Re}(\mathbf{a}|\mathbf{b}) + i \operatorname{Im}(\mathbf{a}|\mathbf{b})$ 和 $\operatorname{Im}(\alpha) = \operatorname{Re}(-i\alpha) \forall \alpha \in \mathbb{F}$, 有 $\operatorname{Im}(\mathbf{a}|\mathbf{b}) = \operatorname{Re}(-i(\mathbf{a}|\mathbf{b})) = \operatorname{Re}(\mathbf{a}|i\mathbf{b})$, 故 $(\mathbf{a}|\mathbf{b}) = \operatorname{Re}(\mathbf{a}|\mathbf{b}) + i \operatorname{Re}(\mathbf{a}|i\mathbf{b})$ 。□

除了内积空间, 我们还可以为一个向量空间引入范的规定, 得到赋范向量空间。

定义 0.9 (向量的范). 设 \mathcal{V} 是数域 \mathbb{F} 向量空间, \mathcal{V} 上的范 (norm) 作用于 \mathcal{V} 中的任一向量, 记为 $\|\mathbf{x}\|$, 并满足:

1. 非负性: $\|\mathbf{x}\| \geq 0, \forall \mathbf{x} \in \mathcal{V}$
2. 调和性: $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|, \forall \mathbf{x} \in \mathcal{V}, \alpha \in \mathbb{F}$
3. 三角不等式: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \forall \mathbf{x}, \mathbf{y} \in \mathcal{V}$

带有一种范的定义的向量空间叫赋范向量空间 (normed vector space)。

我们注意到，赋范向量空间也有一个总成立的不等式——三角不等式。与内积空间的柯西-施瓦茨不等式不同，赋范向量空间的三角不等式是在范的定义中直接规定的，而无法作为定理由之前两个规定证明出来。我们主观上就希望向量的范满足这样的性质，因为“范”是我们为向量定义的一种“长度”的概念，所以希望它能与欧几里得几何公设规定的性质相兼容。

一个赋范向量空间 \mathcal{V} 中，范为 1 的向量称为单位向量 (unit vector)，在本讲义中单位向量会加一个小帽子来特别表示： $\|\hat{\mathbf{a}}\| = 1, \hat{\mathbf{a}} \in \mathcal{V}$ 。赋范向量空间 \mathcal{V} 的任一向量 \mathbf{x} 均可通过 $\hat{\mathbf{x}} = \mathbf{x} / \|\mathbf{x}\|$ 归一化为一个单位向量（由范的调和性易验）。

不管是内积的定义、还是范的定义，都没有具体规定计算方法。只要满足定义的一般要求，我们可以为同一个向量空间赋予不同内积或范的规定。向量的范的其中一种常用的定义是：设 \mathcal{V} 是内积空间，向量 $\mathbf{a} \in \mathcal{V}$ 的范 $\|\mathbf{a}\| \equiv (\mathbf{a}|\mathbf{a})^{\frac{1}{2}}$ 。我们把这一定义称为欧几里得范 (Euclidean norm)。其他范的定义则为非欧几里得范，例如在 \mathbb{R}^n 上，还可以有如下范的定义。对任一 $\mathbf{x} = (x_1, \dots, x_n)$ ， $\|\mathbf{x}\| = \max\{|x_1|, \dots, |x_n|\}^*$ 。一般地，由于给一个向量空间引入内积定义的方式本身就可以有多种，依赖内积的范的定义也会有多种。对于有些向量空间，范的定义可以不依赖内积。

由于有限维向量空间上定义的不同范之间是等价的（见 §??），故我们实际只需以欧几里得范为代表进行后续的讨论，未经说明的话，一般提到有限维向量空间上的范都指欧几里得范。

以下定理说明，在一定条件下，我们能够用范的一般定义构造一个内积，使任何一个尚未定义内积的赋范空间成为一个内积空间。而且特别地，这一种范就是欧几里得范。

定理 0.3. 一个赋范向量空间是一个内积空间，当且仅当该空间的范满足极化恒等式 (polarization identity)，即

$$\|\mathbf{a} + \mathbf{b}\|^2 + \|\mathbf{a} - \mathbf{b}\|^2 = 2\|\mathbf{a}\|^2 + 2\|\mathbf{b}\|^2$$

证明. 设 \mathcal{V} 是数域 \mathbb{F} 上的一个赋范向量空间，若定义二元运算：

$$\begin{aligned} (\mathbf{a}|\mathbf{b}) &= \frac{1}{4}\|\mathbf{a} + \mathbf{b}\|^2 - \frac{1}{4}\|\mathbf{a} - \mathbf{b}\|^2 + \frac{i}{4}\|\mathbf{a} + i\mathbf{b}\|^2 - \frac{i}{4}\|\mathbf{a} - i\mathbf{b}\|^2 \\ &= \frac{1}{4} \sum_{n=1}^4 i^n \|\mathbf{a} + i^n \mathbf{b}\|^2, \forall \mathbf{a}, \mathbf{b} \in \mathcal{V} \end{aligned}$$

可验证上式满足内积定义，且 $\|\mathbf{a}\|^2 = (\mathbf{a}|\mathbf{a})^{\frac{1}{2}}$ ，即该范就是欧几里得范。□

上面的等式在几何上等价于平行四边形法则 (parallelogram law)。特别地，对于 $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ 有 $\mathbf{a} \cdot \mathbf{b} = \frac{1}{4}\|\mathbf{a} + \mathbf{b}\|^2 - \frac{1}{4}\|\mathbf{a} - \mathbf{b}\|^2$ 。

*请验证它满足定义0.9。

柯西-施瓦茨不等式、三角不等式和极化恒等式在很多定理的证明中经常用到，但它们的含义及适用范围需要区分清楚。由定理0.2，柯西-施瓦茨不等式是对任一内积空间均成立的。由定义0.9，三角不等式是对任一赋范向量空间均成立的。而由定理0.3可知，满足极化恒等式的赋范空间必然也是一个内积空间，从而两种空间得到了统一。特别地，欧几里得范是同时满足三角不等式和极化恒等式的范。

下面我们由内积空间的性质引入正交（orthogonal）及相关的概念。

定义 0.10 (正交). 给定内积空间 \mathcal{V} 中的向量 $\mathbf{a}, \mathbf{b} \in \mathcal{V}$ ，若 $(\mathbf{a}|\mathbf{b}) = 0$ ，则称 \mathbf{a} 与 \mathbf{b} 是正交的（orthogonal）。若 S 是 \mathcal{V} 的一个子集，且 S 中的向量两两正交，则称 S 是 \mathcal{V} 的一个正交集（orthogonal set）。若 \mathcal{V} 还是一个赋范向量空间，且 \mathcal{V} 的一个正交集 S 中的向量均满足 $\|\hat{\mathbf{e}}\| = 1 \forall \hat{\mathbf{e}} \in S$ ，则称 S 是规范正交集（orthonormal set）。

易验，零向量与同一内积空间的任意向量都正交。

例 0.6. [1]§7.2 例题 2.2 在数域 \mathbb{F} 上的 $n \times n$ 矩阵的空间 $\mathbb{F}^{n \times n}$ 中，记 E^{pq} 为仅第 p 行、第 q 列为 1，其余为 0 的 $n \times n$ 矩阵。则由这 n^2 个矩阵 $E^{pq}, p = 1, \dots, n, q = 1, \dots, n$ 组成的集合是规范正交集。其中 $\mathbb{F}^{n \times n}$ 上的内积定义是 $(A|B) \equiv \sum_{j,k} A_{jk} \overline{B_{jk}}, \forall A, B \in \mathbb{F}^{n \times n}$ 。

以下定理证明任意一组两两正交的非零向量线性无关。限定“非零”是因为，含零向量的任何向量组线性相关，但零向量又与任一向量正交。还需注意的是线性无关向量组未必都是两两正交的。

定理 0.4. 正交集中的所有非零向量线性无关。

证明. 设 \mathcal{V} 是内积空间， S 是 \mathcal{V} 的一个正交集， $\mathbf{a}_1, \dots, \mathbf{a}_m \in S$ 。令 $\mathbf{b} = \beta_1 \mathbf{a}_1 + \dots + \beta_m \mathbf{a}_m$ ，则

$$\begin{aligned} (\mathbf{b}|\mathbf{a}_k) &= \left(\sum_j \beta_j \mathbf{a}_j \middle| \mathbf{a}_k \right) \\ &= \sum_j \beta_j (\mathbf{a}_j|\mathbf{a}_k) \\ &= \beta_k (\mathbf{a}_k|\mathbf{a}_k), k = 1, \dots, m \\ \because (\mathbf{a}_k|\mathbf{a}_k) &\neq 0 \\ \therefore \beta_k &= \frac{(\mathbf{b}|\mathbf{a}_k)}{\|\mathbf{a}_k\|^2}, k = 1, \dots, m \end{aligned}$$

考察上式可验证 $\mathbf{b} = \mathbf{0} \Leftrightarrow \beta_1 = \dots = \beta_m = 0$ 。

□

由定理0.4以及子空间、线性生成空间的定义 (0.10、0.4)，内积空间 \mathcal{V} 的任一正交集 S 总能线性生成 \mathcal{V} 的一个子空间。若内积空间 \mathcal{V} 的一组基 B 是正交集，则称 B 为 \mathcal{V} 的正交基 (orthogonal basis)。如果 \mathcal{V} 是赋范内积空间，其一组基 B 是规范正交集，则称 B 是 \mathcal{V} 的一组规范正交基 (orthonormal basis)。

定理0.4的证明也给出了由一个赋范内积空间的任一组基经格拉姆-施密特正交化过程 (Gram-Schmidt process) 获得一组规范正交基的方法，作为定理如下。

定理 0.5. 设 \mathcal{V} 是内积空间， $\mathbf{b}_1, \dots, \mathbf{b}_n \in \mathcal{V}$ 是一组线性无关向量。那么总是可以由它们构建一组两两正交的向量 $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathcal{V}$ 使得对于每一 $k = 1, \dots, n$ ，向量组 $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ 都是由 $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ 线性生成的空间的一组基。

证明. 采用数学归纳法。作为 $k = 1$ 的情况，令 $\mathbf{a}_1 = \mathbf{b}_1$ ，则命题显然成立。假设当 $k = m$ 时命题成立，即 $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}, m < n$ 是已经构建好的满足命题要求的正交向量，则对每一 $k = 1, \dots, m$ ， $\{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ 是由 $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ 线性生成的子空间的正交基。令

$$\mathbf{a}_{m+1} = \mathbf{b}_{m+1} - \sum_{k=1}^m \frac{(\mathbf{b}_{m+1} | \mathbf{a}_k)}{(\mathbf{a}_k | \mathbf{a}_k)} \mathbf{a}_k$$

则有 $\mathbf{a}_{m+1} \neq \mathbf{0}$ ，否则 $\mathbf{b}_1, \dots, \mathbf{b}_m, \mathbf{b}_{m+1}$ 线性相关，因 \mathbf{b}_{m+1} 可由 $\mathbf{b}_1, \dots, \mathbf{b}_m$ 线性表出。由上式还可知，对每一 $j = 1, \dots, m$ 均有

$$\begin{aligned} (\mathbf{a}_{m+1} | \mathbf{a}_j) &= (\mathbf{b}_{m+1} | \mathbf{a}_j) - \sum_{k=1}^m \frac{(\mathbf{b}_{m+1} | \mathbf{a}_k)}{(\mathbf{a}_k | \mathbf{a}_k)} (\mathbf{a}_k | \mathbf{a}_j) \\ &= (\mathbf{b}_{m+1} | \mathbf{a}_j) - (\mathbf{b}_{m+1} | \mathbf{a}_j) \\ &= 0 \end{aligned}$$

所以 $\{\mathbf{a}_1, \dots, \mathbf{a}_{m+1}\}$ 是一个非零正交集。由定理0.4，它们线性无关。故 $\{\mathbf{a}_1, \dots, \mathbf{a}_{m+1}\}$ 也是由 $\{\mathbf{b}_1, \dots, \mathbf{b}_{m+1}\}$ 线性生成的子空间的正交基。□

特别地，对 $n = 4$,

$$\begin{aligned} \mathbf{a}_1 &= \mathbf{b}_1 \\ \mathbf{a}_2 &= \mathbf{b}_2 - \frac{(\mathbf{b}_2 | \mathbf{a}_1)}{(\mathbf{a}_1 | \mathbf{a}_1)} \mathbf{a}_1 \\ \mathbf{a}_3 &= \mathbf{b}_3 - \frac{(\mathbf{b}_3 | \mathbf{a}_1)}{(\mathbf{a}_1 | \mathbf{a}_1)} \mathbf{a}_1 - \frac{(\mathbf{b}_3 | \mathbf{a}_2)}{(\mathbf{a}_2 | \mathbf{a}_2)} \mathbf{a}_2 \\ \mathbf{a}_4 &= \mathbf{b}_4 - \frac{(\mathbf{b}_4 | \mathbf{a}_1)}{(\mathbf{a}_1 | \mathbf{a}_1)} \mathbf{a}_1 - \frac{(\mathbf{b}_4 | \mathbf{a}_2)}{(\mathbf{a}_2 | \mathbf{a}_2)} \mathbf{a}_2 - \frac{(\mathbf{b}_4 | \mathbf{a}_3)}{(\mathbf{a}_3 | \mathbf{a}_3)} \mathbf{a}_3 \end{aligned}$$

推论 0.5.1. 每个有限维赋范内积空间都有一组规范正交基。

证明. 只需要对采用格拉姆-施密特正交化过程得到的正交基, 再对其基向量归一化即可。□

例 0.7. 考虑 \mathbb{R}^3 中的三个向量 $\mathbf{b}_1 = (3, 0, 4)^\top$, $\mathbf{b}_2 = (-1, 0, 7)^\top$, $\mathbf{b}_3 = (2, 9, 11)^\top$ 。首先可以验证它们线性无关, 即

$$\begin{cases} 3x_1 - x_2 + 2x_3 = 0 \\ 9x_3 = 0 \\ 4x_1 + 7x_2 + 11x_3 = 0 \end{cases}$$

只有唯一解 $x_1 = x_2 = x_3 = 0$ 。

通过格拉姆-施密特正交化过程可由 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ 得到 \mathbb{R}^3 的一组正交基:

$$\begin{aligned} \mathbf{a}_1 &= (3, 0, 4)^\top \\ \mathbf{a}_2 &= (-1, 0, 7)^\top - \frac{(-1, 0, 7)^\top \cdot (3, 0, 4)^\top}{(3, 0, 4)^\top \cdot (3, 0, 4)^\top} (3, 0, 4)^\top \\ &= (-1, 0, 7)^\top - (3, 0, 4)^\top \\ &= (-4, 0, 3)^\top \\ \mathbf{a}_3 &= (2, 9, 11)^\top - \frac{(2, 9, 11)^\top \cdot (3, 0, 4)^\top}{(3, 0, 4)^\top \cdot (3, 0, 4)^\top} (3, 0, 4)^\top - \frac{(2, 9, 11)^\top \cdot (-4, 0, 3)^\top}{(-4, 0, 3)^\top \cdot (-4, 0, 3)^\top} (-4, 0, 3)^\top \\ &= (2, 9, 11)^\top - 2(3, 0, 4)^\top - (-4, 0, 3)^\top \\ &= (0, 9, 0)^\top \end{aligned}$$

可见 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 均为非零向量, 故它们是 \mathbb{R}^3 的一组正交基。归一化后得到 $\hat{\mathbf{a}}_1 = \frac{1}{5}\mathbf{a}_1, \hat{\mathbf{a}}_2 = \frac{1}{5}\mathbf{a}_2, \hat{\mathbf{a}}_3 = (0, 1, 0)^\top$ 是一组规范正交基。任一向量 $\mathbf{x} = (x_1, x_2, x_3)^\top \in \mathbb{R}^3$ 在基 $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ 下的坐标为:

$$(x_1, x_2, x_3)^\top = \frac{3x_1 + 4x_3}{25}\mathbf{a}_1 + \frac{-4x_1 + 3x_3}{25}\mathbf{a}_2 + \frac{x_2}{9}\mathbf{a}_3$$

由格拉姆-施密特正交化过程可知, 一般地, 对于规范正交基 $\{\mathbf{a}_i\}$ 有

$$\mathbf{a}_i \cdot \mathbf{a}_j = \delta_{ij}$$

其中

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

叫克劳内克符号 (Kronecker symbol)。

在选取什么基之下, 向量 $\mathbf{x} = (x_1, \dots, x_n)$ 在该基下的坐标就恰好是 x_1, \dots, x_n 呢? 这样的基 $\{\mathbf{e}_i\}$ 叫标准基 (standard basis), 其中

$$\hat{\mathbf{e}}_i = (e_{i,1}, \dots, e_{i,n})^\top, e_{i,j} = \delta_{ij}$$

标准基是一个规范正交基。

以下推导向量内积在给定基下的坐标计算公式。设 \mathcal{V} 是数域 \mathbb{F} 上的 n 维内积空间， $\{\mathbf{e}_i\}$ 是 \mathcal{V} 的一组基，则任意两个向量 \mathbf{a}, \mathbf{b} 可表示为

$$\mathbf{a} = \sum_{i=1}^n \alpha_i \mathbf{e}_i, \mathbf{b} = \sum_{i=1}^n \beta_i \mathbf{e}_i$$

它们的内积

$$\begin{aligned} (\mathbf{a}|\mathbf{b}) &= \left(\sum_{i=1}^n \alpha_i \mathbf{e}_i \left| \sum_{j=1}^n \beta_j \mathbf{e}_j \right. \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\beta_j} (\mathbf{e}_i|\mathbf{e}_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\beta_j} G_{ij} \end{aligned}$$

其中， G 为基 $\{\mathbf{e}_i\}$ 的格拉姆矩阵 (Gram matrix)，即 $G_{ij} = (\mathbf{e}_i|\mathbf{e}_j)$ ，由内积定义有 $G_{ij} = \overline{G_{ji}}$ 。若 $\{\mathbf{e}_i\}$ 是正交基，则 $G_{ij} = \overline{G_{ji}}\delta_{ij}$ 。若 $\{\mathbf{e}_i\}$ 是规范正交基，则 $G_{ij} = \delta_{ij}$ 。