欧几里得空间

在熟知的三维空间 $\mathbb{R}^3$ 中, 我们可以建立直角坐标系 $Oxyz$. 对于任意的向量 $\boldsymbol{a} = (a_1, a_2, a_3)^{\mathsf{T}}$ 和 $\boldsymbol{b} = (b_1, b_2, b_3)^{\mathsf{T}} \in \mathbb{R}^3$, 我们有向量的内积 (点乘, 标量积)

$$\boldsymbol{a} \cdot \boldsymbol{b} := a_1 b_1 + a_2 b_2 + a_3 b_3.$$

这儿, 我们将把内积记成 $(\boldsymbol{a}, \boldsymbol{b})$. 内积的几何意义是: 向量 $\boldsymbol{a}$ 的模长 $|\boldsymbol{a}|$ 为 $\sqrt{(\boldsymbol{a}, \boldsymbol{a})}$, 而且

$$(\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \cdot \cos(\theta),$$

其中的 θ 是向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 之间的夹角. 这一套处理方法可以推广到高维情形.

定义 1

设 V 是实数域 $\mathbb{R}$ 上的线性空间. 如果对于 V 内任意两个向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 都按照某一法则对应于一个实数, 记作 $(\boldsymbol{a}, \boldsymbol{b}) \in \mathbb{R}$,[1] 满足:

- (对称性) $(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{a})$,
- (对于第一位置的线性性) $\lambda(\boldsymbol{a}, \boldsymbol{b}) = (\lambda\boldsymbol{a}, \boldsymbol{b})$ 以及 $(\boldsymbol{a}+\boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a}, \boldsymbol{c}) + (\boldsymbol{b}, \boldsymbol{c})$,
- (正定性) 对于任意的 $\boldsymbol{a} \in V$ 有 $(\boldsymbol{a}, \boldsymbol{a}) \geq 0$, 且不等式中等号成立的充要条件为 $\boldsymbol{a} = \boldsymbol{0}$,

则称该二元运算法则为 V 上的内积 (*inner product*), 而 $(\boldsymbol{a}, \boldsymbol{b})$ 称为 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的内积. 定义了内积的 $\mathbb{R}$ 上的线性空间称为内积空间 (*inner product space*) 或欧几里得空间, 简称为欧氏空间. 此时, 把线性空间 V 的维数叫作欧氏空间 V 的维数.

[1] 有些教材里将其记作 $\langle \boldsymbol{a}, \boldsymbol{b} \rangle$, 用来与有序对的记号作区别; 这样的记号也许更为合理.

注 2

1. 从内积的对称性和对于第一位置的线性性，我们可以马上推出对于第二位置的线性性.
2. 对于任意的 $\boldsymbol{a} \in V$，我们有 $(\boldsymbol{a}, \mathbf{0}) = (\mathbf{0}, \boldsymbol{a}) = 0$.
3. 同一个线性空间在不同的内积下将被视作不同的内积空间.

内积空间的例子

1. 对于数组空间 $\mathbb{R}^n$, 将里面的向量视为列向量. 任取其中两个向量 $\boldsymbol{a}=(a_1,\ldots,a_n)^{\mathsf{T}}$, $\boldsymbol{b}=(b_1,\ldots,b_n)^{\mathsf{T}}$, 我们可以定义

$$(\boldsymbol{a},\boldsymbol{b}) := a_1b_1 + a_2b_2 + \cdots + a_nb_n = \boldsymbol{a}^{\mathsf{T}}\boldsymbol{b}.$$

可以验证, 这是一个内积, 称为 $\mathbb{R}^n$ 的标准内积 (standard inner product) 或者欧几里得内积. 若不作特别说明的话, 我们默认 $\mathbb{R}^n$ 为采用了如此定义的内积空间. 当然, 对于给定的一个元素全为正的向量 $\boldsymbol{w}=(w_1,\ldots,w_n)^{\mathsf{T}}$, 我们也可以定义 $\mathbb{R}^n$ 上的一个新的内积为

$$(\boldsymbol{a},\boldsymbol{b}) := \sum_{i=1}^{n} a_ib_iw_i.$$

其中这些标量 w_i 称为这个新的内积的权 (weights).

内积空间的例子

2. 设 V 是 $\mathbb{R}$ 上的 n 维线性空间, $\alpha_1, \ldots, \alpha_n$ 为 V 的一组基. 对于任意的 $\boldsymbol{x}, \boldsymbol{y} \in V$, 设它们在 $\alpha_1, \ldots, \alpha_n$ 下的坐标分别为 $\boldsymbol{X} = (x_1, \ldots, x_n)^\mathsf{T}$ 和 $\boldsymbol{Y} = (y_1, \ldots, y_n)^\mathsf{T}$, 即 $\boldsymbol{x} = \sum_i x_i \alpha_i$, $\boldsymbol{y} = \sum_i y_i \alpha_i$. 定义

$$(\boldsymbol{x}, \boldsymbol{y}) := \sum_i x_i y_i,$$

即 $\mathbb{R}^n$ 中的向量 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 之间的标准内积. 不难验证, 这给出了 V 上的一个内积.

3. 由全体 $m \times n$ 实矩阵构成的线性空间 $\mathbb{R}^{m\times n}$ 中, 定义一个二元函数为

$$(\boldsymbol{A}, \boldsymbol{B}) := \mathrm{tr}(\boldsymbol{A}^\mathsf{T} \boldsymbol{B}).$$

容易验证, 该二元函数是 $\mathbb{R}^{m\times n}$ 上的一个内积.

内积空间的例子

4. $C[a,b]$ 为有界闭区间 $[a,b]$ 上的实连续函数的全体, 是 $\mathbb{R}$ 上的无穷维线性空间. 这儿假定 $a<b$. 对于 $f,g\in C[a,b]$, 定义

$$(f,g):=\int_a^b f(x)g(x)\,\mathrm{d}x.$$

这给出了 $C[a,b]$ 的一个内积. 为了验证正定性, 我们注意到

$$(f,f)=\int_a^b f(x)^2\,\mathrm{d}x\geq 0.$$

若存在 $[a,b]$ 中某个 x_0 满足 $f(x_0)\neq 0$, 那么由函数 f^2 的连续性可知, 存在 $[a,b]$ 中包含 x_0 的小区间 I, 使得在 I 上有 $f^2\geq\frac{f(x_0)^2}{2}$.

内积空间的例子

此时，若令 I 的长度为 $p>0$，那么

$$(f,f)=\int_a^b f(x)^2\,\mathrm{d}x \ge \int_I f(x)^2\,\mathrm{d}x \ge \frac{f(x_0)^2 p}{2} > 0.$$

这说明若 $(f,f)=0$，则 f 在 $[a,b]$ 上必为零函数.

更一般地，若 $w(x)\in C[a,b]$ 是一个正的连续函数，那么

$$(f,g):=\int_a^b f(x)g(x)w(x)\,\mathrm{d}x$$

给出了 $C[a,b]$ 的一个新的内积，而 $w(x)$ 称为该内积的权函数 (weight function).

内积空间的例子

5. 对于 $F=\mathbb{R}$, 我们考虑线性空间 $F_n[x]$. 这是次数不超过 n 的实系数多项式构成的实线性空间. 设 $a_0, a_1, \ldots, a_n$ 为不同的实数, 对于 $f, g \in F_n[x]$, 定义

$$(f, g) := \sum_{i=0}^{n} f(a_i) g(a_i).$$

这给出了 $F_n[x]$ 的一个内积. 为了验证正定性, 我们注意到

$$(f, f) = \sum_{i=0}^{n} f(a_i)^2 \geq 0.$$

若 $(f, f) = 0$, 则 $a_0, a_1, \ldots, a_n$ 为 $f(x)$ 的不同的根. 而 $f(x)$ 的次数不超过 n, 这说明 f 必为零多项式.

内积空间的例子

更一般地，若 $w_0, w_1, \ldots, w_n$ 为一组正数，那么

$$(f, g) := \sum_{i=0}^{n} f(a_i) g(a_i) w_i$$

也定义了 $F_n[x]$ 上的一个内积.

定理 3 (Cauchy–Schwarz)

设 V 是欧氏空间, $(\cdot,\cdot)$ 是 V 的内积, 则对 V 中的任意两个向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$, 有 $|(\boldsymbol{a},\boldsymbol{b})| \le \sqrt{(\boldsymbol{a},\boldsymbol{a})(\boldsymbol{b},\boldsymbol{b})}$.

注 4

1. 由证明过程易知，*Cauchy–Schwarz* 不等式中等号成立的条件是向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 线性相关.
2. 将 *Cauchy–Schwarz* 不等式运用于前面的第 *1* 个例子，我们得到，若 $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$，则

$$|a_1 b_1 + a_2 b_2 + \cdots + a_n b_n| \le \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2}.$$

3. 将 *Cauchy–Schwarz* 不等式运用于前面的第 *4* 个例子，我们得到，若 $f(x), g(x) \in C[a, b]$，则

$$\left|\int_a^b f(x)g(x)\,\mathrm{d}x\right| \le \sqrt{\int_a^b f(x)^2\,\mathrm{d}x}\sqrt{\int_a^b g(x)^2\,\mathrm{d}x}.$$

定义 5

设 V 是欧氏空间, $(\cdot,\cdot)$ 是 V 的内积, $\boldsymbol{a}, \boldsymbol{b} \in V$. 仿照三维的情形, 我们定义如下.

1. $|\boldsymbol{a}| := \sqrt{(\boldsymbol{a}, \boldsymbol{a})} \geq 0$ 称为 $\boldsymbol{a}$ 的长度 (*length*) 或模 (*module*). 有时模也记作 $\|\boldsymbol{a}\|$.[2] 由内积的正定性容易推出, $\|\boldsymbol{a}\| = 0$ 当且仅当 $\boldsymbol{a} = \mathbf{0}$. 另外, 对任意的 $\lambda \in \mathbb{R}$, 我们都有 $\|\lambda \boldsymbol{a}\| = |\lambda| \cdot \|\boldsymbol{a}\|$.
2. 两个非零向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 之间的夹角为 $\theta = \arccos\left(\frac{(\boldsymbol{a},\boldsymbol{b})}{\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|}\right)$. 注意, 由 *Cauchy–Schwarz* 不等式, 我们有

$$-1 \leq \frac{(\boldsymbol{a}, \boldsymbol{b})}{\|\boldsymbol{a}\| \cdot \|\boldsymbol{b}\|} \leq 1,$$

从而定义是可行的.

[2] 为了与标量的绝对值以示区别, 在我们的教案里会采用这一记号.

3. 进一步地, 若 $(\boldsymbol{a},\boldsymbol{b})=0$ (允许 $\boldsymbol{a},\boldsymbol{b}$ 为 $\mathbf{0}$), 则称 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 垂直 (perpendicular) 或正交 (orthogonal), 记作 $\boldsymbol{a}\perp\boldsymbol{b}$. 容易验证, $\boldsymbol{a}\perp\boldsymbol{b}$ 当且仅当 $\boldsymbol{b}\perp\boldsymbol{a}$, 当且仅当勾股定理 (也被称为毕达哥拉斯定理) 成立:

$$\|\boldsymbol{a}+\boldsymbol{b}\|^2=\|\boldsymbol{a}\|^2+\|\boldsymbol{b}\|^2.$$

另外, 零向量 $\mathbf{0}$ 与 V 中任意的向量都正交, 并且是 V 中满足这一条件的唯一的向量, 也是 V 中唯一的与自身正交的向量.

4. 长度为 1 的向量称为单位向量.

5. 向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 之间的距离定义为 $d(\boldsymbol{\alpha},\boldsymbol{\beta}) := \|\boldsymbol{\alpha}-\boldsymbol{\beta}\|$. 我们很容易验证:

(对称性) $d(\boldsymbol{\alpha},\boldsymbol{\beta}) = d(\boldsymbol{\beta},\boldsymbol{\alpha})$,

(正定性) $d(\boldsymbol{\alpha},\boldsymbol{\beta}) \geq 0$, 且 $d(\boldsymbol{\alpha},\boldsymbol{\beta}) = 0$ 当且仅当 $\boldsymbol{\alpha}=\boldsymbol{\beta}$,

(三角不等式) $d(\boldsymbol{\alpha},\boldsymbol{\beta}) \leq d(\boldsymbol{\alpha},\boldsymbol{\gamma}) + d(\boldsymbol{\beta},\boldsymbol{\gamma})$.

容易看出,

关于距离的三角不等式

$$\xLeftrightarrow[\boldsymbol{b}=\boldsymbol{\gamma}-\boldsymbol{\beta}]{\text{记 } \boldsymbol{a}=\boldsymbol{\alpha}-\boldsymbol{\gamma}} \|\boldsymbol{a}+\boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\| \quad (\text{模的三角不等式})$$

$$\Leftrightarrow \|\boldsymbol{a}+\boldsymbol{b}\|^2 \leq (\|\boldsymbol{a}\| + \|\boldsymbol{b}\|)^2$$

$$\Leftrightarrow (\boldsymbol{a}+\boldsymbol{b},\boldsymbol{a}+\boldsymbol{b}) \leq \|\boldsymbol{a}\|^2 + 2\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\| + \|\boldsymbol{b}\|^2$$

$$\Leftrightarrow (\boldsymbol{a},\boldsymbol{a}) + (\boldsymbol{a},\boldsymbol{b}) + (\boldsymbol{b},\boldsymbol{a}) + (\boldsymbol{b},\boldsymbol{b}) \leq (\boldsymbol{a},\boldsymbol{a}) + 2\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\| + (\boldsymbol{b},\boldsymbol{b})$$

$$\Leftrightarrow (\boldsymbol{a},\boldsymbol{b}) \leq \|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|. \quad (\text{Cauchy–Schwarz 不等式})$$

内积的表示与标准正交基

回忆

设 V 是有限维向量空间, $\mathscr{A}$ 为 V 上的线性变换, 在基 $\alpha_1, \ldots, \alpha_n$ 下的矩阵为 $\boldsymbol{A}$, 则 $\mathscr{A}$ 的具体作用由 $\boldsymbol{A}$ 唯一决定.

若 V 是有限维欧氏空间, 对于其上的内积, 我们也有类似地表述.

设 $(\cdot,\cdot)$ 是 V 上的内积，任给 $\boldsymbol{x}, \boldsymbol{y} \in V$，假定它们在基 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ 下的坐标分别为 $\boldsymbol{X} = (x_1, \ldots, x_n)^{\mathsf{T}}$ 和 $\boldsymbol{Y} = (y_1, \ldots, y_n)^{\mathsf{T}}$，我们有

$$(\boldsymbol{x}, \boldsymbol{y}) = \left(\sum_{i=1}^{n} x_i \boldsymbol{\alpha}_i, \sum_{j=1}^{n} y_j \boldsymbol{\alpha}_j \right) \xlongequal{\text{双线性性}} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j (\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j).$$

令 $\boldsymbol{G} = (g_{ij}) \in \mathbb{R}^{n \times n}$，其中 $g_{ij} = (\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j)$. 则上式可以表述成

$$(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{X}^{\mathsf{T}} \boldsymbol{G} \boldsymbol{Y}.$$

注意，由内积的对称性，有 $(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j) = (\boldsymbol{\alpha}_j, \boldsymbol{\alpha}_i)$，即 $g_{ij} = g_{ji}$. 这说明方阵 $\boldsymbol{G}$ 是实对称阵. 在接下来的讨论中，方阵 $\boldsymbol{G}$ 将被称为内积 $(\cdot,\cdot)$ 在基 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ 下的度量矩阵.

定义 6

对于 n 阶实对称阵 $\boldsymbol{A}$, 若是对于任意的非零列向量 $\boldsymbol{X} \in \mathbb{R}^n$, 皆有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X} > 0$, 则称 $\boldsymbol{A}$ 是一个正定方阵.

故由内积的正定性, 知其度量矩阵 $\boldsymbol{G}$ 必为一个正定方阵.

反过来, 给定 $\mathbb{R}$ 上的有限维线性空间 V 和上面的一组基 $\alpha_1, \ldots, \alpha_n$, 以及一个 n 阶正定的实对称方阵 $\boldsymbol{G}$, 可以定义 V 上的一个内积如下. 任给 $\boldsymbol{x}, \boldsymbol{y} \in V$, 假定它们在基 $\alpha_1, \ldots, \alpha_n$ 下的坐标分别为 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$, 定义 $(\boldsymbol{x}, \boldsymbol{y}) := \boldsymbol{X}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{Y}$. 则 V 变成了一个内积空间.

很自然的问题: 在考虑线性空间 V 的一组新的基时, 线性变换相应的矩阵要做相似变换. 那么对于内积而言, 相应的度量矩阵该如何改变?

引理 7

设 $\boldsymbol{A}, \boldsymbol{B} \in F^{n\times n}$. 则 $\boldsymbol{A} = \boldsymbol{B}$ 当且仅当对于任意的 n 维列向量 $\boldsymbol{X}, \boldsymbol{Y}$, 有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Y} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Y}$.

设 $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n$ 是欧氏空间 V 的另外一组基. 设 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 到 $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n$ 的过渡矩阵为 $\boldsymbol{P}$, 即有

$$(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)\boldsymbol{P}.$$

设内积在 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 和 $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n$ 下的度量矩阵分别为 $\boldsymbol{G},\widetilde{\boldsymbol{G}}$. 对于任意的 $\boldsymbol{x},\boldsymbol{y}\in V$, 假设它们在基 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 下的坐标分别为 $\boldsymbol{X},\boldsymbol{Y}$. 从而, 它们在基 $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n$ 下的坐标分别为 $\widetilde{\boldsymbol{X}}=\boldsymbol{P}^{-1}\boldsymbol{X}$ 和 $\widetilde{\boldsymbol{Y}}=\boldsymbol{P}^{-1}\boldsymbol{Y}$. 由定义, 有

$$\boxed{\boldsymbol{X}^{\mathsf{T}}\boldsymbol{G}\boldsymbol{Y}}=(\boldsymbol{x},\boldsymbol{y})=\widetilde{\boldsymbol{X}}^{\mathsf{T}}\widetilde{\boldsymbol{G}}\widetilde{\boldsymbol{Y}}=\boxed{\boldsymbol{X}^{\mathsf{T}}(\boldsymbol{P}^{\mathsf{T}})^{-1}\widetilde{\boldsymbol{G}}\boldsymbol{P}^{-1}\boldsymbol{Y}}.$$

由上面的引理可知，这说明 $\boldsymbol{G} = (\boldsymbol{P}^{\mathsf{T}})^{-1}\widetilde{\boldsymbol{G}}\boldsymbol{P}^{-1}$，即

$$\widetilde{\boldsymbol{G}} = \boldsymbol{P}^{\mathsf{T}}\boldsymbol{G}\boldsymbol{P}.$$

定义 8

给定两个同阶方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$. 若存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{A} = \boldsymbol{P}^{\mathsf{T}}\boldsymbol{B}\boldsymbol{P}$，则称 $\boldsymbol{A}$ 是 $\boldsymbol{B}$ 由相合变换 (congruent transformation) 得到的，称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是相合的.