

内积空间

秦加俊

2026 年 6 月 4 日

1 内积、距离与正交展开

例题 1.1 设 V 是内积空间, $\alpha, \beta \in V$, 证明:

(1) 当 $\mathbb{F} = \mathbb{C}$ 时, 有

$$\operatorname{Im}(\alpha, \beta) = \operatorname{Re}(\alpha, i\beta),$$

且

$$(\alpha, \beta) = \frac{1}{4}\|\alpha + \beta\|^2 - \frac{1}{4}\|\alpha - \beta\|^2 + \frac{i}{4}\|\alpha + i\beta\|^2 - \frac{i}{4}\|\alpha - i\beta\|^2.$$

(2) 当 $\mathbb{F} = \mathbb{R}$ 时, 有

$$(\alpha, \beta) = \frac{1}{4}\|\alpha + \beta\|^2 - \frac{1}{4}\|\alpha - \beta\|^2.$$

于是, 向量的内积由向量的长度唯一确定。

解答. 先讨论复内积空间。设

$$(\alpha, \beta) = a + bi,$$

其中 $a, b \in \mathbb{R}$ 。按照内积关于第二个变量共轭线性的约定, 有

$$(\alpha, i\beta) = \bar{i}(\alpha, \beta) = -i(a + bi) = b - ai.$$

因此

$$\operatorname{Re}(\alpha, i\beta) = b = \operatorname{Im}(\alpha, \beta).$$

另一方面,

$$\|\alpha + \beta\|^2 = (\alpha + \beta, \alpha + \beta) = \|\alpha\|^2 + \|\beta\|^2 + (\alpha, \beta) + (\beta, \alpha),$$

而

$$\|\alpha - \beta\|^2 = (\alpha - \beta, \alpha - \beta) = \|\alpha\|^2 + \|\beta\|^2 - (\alpha, \beta) - (\beta, \alpha).$$

两式相减得

$$\|\alpha + \beta\|^2 - \|\alpha - \beta\|^2 = 2((\alpha, \beta) + (\beta, \alpha)) = 4\operatorname{Re}(\alpha, \beta).$$

因此

$$\operatorname{Re}(\alpha, \beta) = \frac{1}{4}\|\alpha + \beta\|^2 - \frac{1}{4}\|\alpha - \beta\|^2.$$

同理,

$$\operatorname{Im}(\alpha, \beta) = \operatorname{Re}(\alpha, i\beta) = \frac{1}{4}\|\alpha + i\beta\|^2 - \frac{1}{4}\|\alpha - i\beta\|^2.$$

于是

$$(\alpha, \beta) = \operatorname{Re}(\alpha, \beta) + i \operatorname{Im}(\alpha, \beta),$$

即

$$(\alpha, \beta) = \frac{1}{4}\|\alpha + \beta\|^2 - \frac{1}{4}\|\alpha - \beta\|^2 + \frac{i}{4}\|\alpha + i\beta\|^2 - \frac{i}{4}\|\alpha - i\beta\|^2.$$

当 $\mathbb{F} = \mathbb{R}$ 时, 内积本身为实数, 故由上面的实部公式直接得到

$$(\alpha, \beta) = \frac{1}{4}\|\alpha + \beta\|^2 - \frac{1}{4}\|\alpha - \beta\|^2.$$

这说明只要知道所有向量的长度, 就可以由极化恒等式恢复任意两向量的内积。□

注记 1.1

在赋范线性空间中, 范数一般只刻画向量的长度, 而不一定能够刻画向量之间的角度。因此, 并不是每一个范数都可以由某个内积诱导出来。

一个重要判别准则是 Jordan-von Neumann 定理: 设 X 是实或复赋范线性空间, 则 X 的范数来自某个内积, 当且仅当该范数满足平行四边形恒等式

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2, \quad x, y \in X.$$

也就是说, 平行四边形恒等式正是范数能够由内积诱导的充要条件。

在实赋范空间中, 如果范数满足平行四边形恒等式, 则可以通过极化恒等式定义内积:

$$(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2).$$

在复赋范空间中, 相应的极化恒等式为

$$(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2).$$

由此定义出的内积满足

$$(x, x) = \|x\|^2,$$

从而它诱导出的范数正是原来的范数。

例题 1.2 设 X 是非空集合, 如果映射 $d: X \times X \rightarrow \mathbb{R}$ 满足:

(i) 正定性: 对任意 $x, y \in X$, 有 $d(x, y) \geq 0$, 且当且仅当 $x = y$ 时等号成立;

(ii) 对称性: 对任意 $x, y \in X$, 有

$$d(x, y) = d(y, x);$$

(iii) 三角形不等式: 对任意 $x, y, z \in X$, 有

$$d(x, z) \leq d(x, y) + d(y, z),$$

则称 d 是一个距离, 称 $d(x, y)$ 为 x 和 y 之间的距离。此时称 X 是一个度量空间。

设 V 是内积空间, 对任意 $\alpha, \beta \in V$, 定义

$$d(\alpha, \beta) = \|\alpha - \beta\|.$$

证明： d 是一个距离，从而 V 在该距离下成为度量空间。

解答. 依次验证距离的三个条件。

首先，由范数的非负性，

$$d(\alpha, \beta) = \|\alpha - \beta\| \geq 0.$$

且

$$d(\alpha, \beta) = 0 \iff \|\alpha - \beta\| = 0 \iff \alpha - \beta = 0 \iff \alpha = \beta.$$

故正定性成立。

其次，

$$d(\alpha, \beta) = \|\alpha - \beta\| = \|-(\beta - \alpha)\| = \|\beta - \alpha\| = d(\beta, \alpha),$$

故对称性成立。

最后，对任意 $\alpha, \beta, \gamma \in V$ ，由范数的三角形不等式，

$$d(\alpha, \gamma) = \|\alpha - \gamma\| = \|(\alpha - \beta) + (\beta - \gamma)\| \leq \|\alpha - \beta\| + \|\beta - \gamma\|.$$

即

$$d(\alpha, \gamma) \leq d(\alpha, \beta) + d(\beta, \gamma).$$

故三角形不等式成立。

综上， d 是 V 上的一个距离，因此 V 在该距离下成为度量空间。 $\square$

知识 1.1: 从内积空间到拓扑空间

内积空间、赋范线性空间、度量空间和拓扑空间之间有一条自然的“遗忘结构”的链条：

$$\text{内积空间} \implies \text{赋范线性空间} \implies \text{度量空间} \implies \text{拓扑空间}.$$

也就是说，每往右走一步，空间结构更弱，能够讨论的几何性质也更少。

空间类型	基本结构	诱导关系	能讨论的典型概念
内积空间	给定内积 (x, y)	内积诱导范数： $\ x\ = \sqrt{(x, x)}$	长度、角度、正交、投影、Fourier 展开、Parseval 等式
赋范线性空间	给定范数 $\ x\ $	范数诱导距离： $d(x, y) = \ x - y\ $	长度、收敛、连续性、完备性、级数收敛
度量空间	给定距离函数 $d(x, y)$	距离诱导开集： $U \text{ 开} \iff \forall x \in U, \exists r > 0, B(x, r) \subset U$	距离、极限、连续性、紧性、完备性、稠密性
拓扑空间	给定开集族 $\mathcal{T}$	直接用开集描述空间结构	开集、闭集、邻域、连续映射、连通性、紧性

其中每一步的具体诱导方式为：

$$(x, y) \implies \|x\| = \sqrt{(x, x)} \implies d(x, y) = \|x - y\| \implies \mathcal{T}_d.$$

这里 $\mathcal{T}_d$ 表示由度量 d 诱导出的拓扑，即由所有开球

$$B(x, r) = \{y \in X : d(x, y) < r\}$$

生成的开集族。
需要注意的是，反过来一般不成立。也就是说，一个拓扑空间不一定来自某个度量；一个度量空间不一定来自某个范数；一个赋范线性空间的范数也不一定来自某个内积。赋范线性空间的范数能够由内积诱导，当且仅当它满足平行四边形恒等式：

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

例题 1.3 Bessel 不等式 设 $\alpha_1, \alpha_2, \dots, \alpha_m$ 是 n 维内积空间 V 的一组两两正交的单位向量, $\alpha \in V$ 。证明:

$$\sum_{i=1}^m |(\alpha, \alpha_i)|^2 \leq \|\alpha\|^2,$$

且向量

$$\beta = \alpha - \sum_{i=1}^m (\alpha, \alpha_i) \alpha_i$$

与 $\alpha_1, \alpha_2, \dots, \alpha_m$ 都正交。

解答. 先证明正交性。对任意 $1 \leq j \leq m$, 有

$$\begin{aligned} (\beta, \alpha_j) &= \left(\alpha - \sum_{i=1}^m (\alpha, \alpha_i) \alpha_i, \alpha_j \right) \\ &= (\alpha, \alpha_j) - \sum_{i=1}^m (\alpha, \alpha_i) (\alpha_i, \alpha_j). \end{aligned}$$

由于 $\alpha_1, \alpha_2, \dots, \alpha_m$ 是两两正交的单位向量,

$$(\alpha_i, \alpha_j) = \delta_{ij}.$$

因此

$$(\beta, \alpha_j) = (\alpha, \alpha_j) - (\alpha, \alpha_j) = 0.$$

故 β 与每个 α_j 都正交。

又因为

$$\alpha = \beta + \sum_{i=1}^m (\alpha, \alpha_i) \alpha_i,$$

并且 β 与每个 α_i 正交, 所以由勾股定理,

$$\|\alpha\|^2 = \|\beta\|^2 + \left\| \sum_{i=1}^m (\alpha, \alpha_i) \alpha_i \right\|^2.$$

由于 $\alpha_1, \alpha_2, \dots, \alpha_m$ 是一组两两正交的单位向量,

$$\left\| \sum_{i=1}^m (\alpha, \alpha_i) \alpha_i \right\|^2 = \sum_{i=1}^m |(\alpha, \alpha_i)|^2.$$

于是

$$\|\alpha\|^2 = \|\beta\|^2 + \sum_{i=1}^m |(\alpha, \alpha_i)|^2 \geq \sum_{i=1}^m |(\alpha, \alpha_i)|^2.$$

即得 Bessel 不等式。 □

例题 1.4 设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 n 维内积空间 V 的一组向量。证明：下面命题等价：

(i) $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 V 的标准正交基；

(ii) 对任意 $\alpha \in V$ ，有

$$\alpha = \sum_{i=1}^n (\alpha, \alpha_i) \alpha_i;$$

(iii) **Parseval 等式** 对任意 $\alpha, \beta \in V$ ，有

$$(\alpha, \beta) = \sum_{i=1}^n (\alpha, \alpha_i)(\alpha_i, \beta);$$

(iv) 对任意 $\alpha \in V$ ，有

$$\|\alpha\|^2 = \sum_{i=1}^n |(\alpha, \alpha_i)|^2.$$

解答. 我们证明

$$(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i).$$

$(i) \Rightarrow (ii)$: 若 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 V 的标准正交基，则任意 $\alpha \in V$ 可以唯一表示为

$$\alpha = \sum_{i=1}^n c_i \alpha_i.$$

两边与 α_j 作内积，得到

$$(\alpha, \alpha_j) = \sum_{i=1}^n c_i (\alpha_i, \alpha_j) = c_j.$$

因此

$$\alpha = \sum_{i=1}^n (\alpha, \alpha_i) \alpha_i.$$

$(ii) \Rightarrow (iii)$: 由

$$\alpha = \sum_{i=1}^n (\alpha, \alpha_i) \alpha_i,$$

两边与 β 作内积，得到

$$(\alpha, \beta) = \sum_{i=1}^n (\alpha, \alpha_i)(\alpha_i, \beta).$$

故 Parseval 等式成立。

$(iii) \Rightarrow (iv)$: 在 Parseval 等式中取 $\beta = \alpha$ ，得到

$$\|\alpha\|^2 = (\alpha, \alpha) = \sum_{i=1}^n (\alpha, \alpha_i)(\alpha_i, \alpha) = \sum_{i=1}^n |(\alpha, \alpha_i)|^2.$$

$(iv) \Rightarrow (i)$: 设 $e_1, e_2, \dots, e_n$ 是 V 的一组标准正交基。把每个 α_i 在这组基下的坐标列向量记为 a_i ，并令

$$A = (a_1, a_2, \dots, a_n).$$

也就是说, 矩阵 A 的第 i 列就是 α_i 的坐标。

任取 $\alpha \in V$, 设它在基 $e_1, \dots, e_n$ 下的坐标列向量为 x 。则

$$(\alpha, \alpha_i) = x^* a_i.$$

由条件 (iv), 对任意 x , 有

$$\|x\|^2 = \sum_{i=1}^n |x^* a_i|^2.$$

而

$$\sum_{i=1}^n |x^* a_i|^2 = \|A^* x\|^2.$$

因此

$$\|x\|^2 = \|A^* x\|^2.$$

于是

$$x^* x = x^* A A^* x,$$

即

$$x^* (A A^* - I) x = 0$$

对任意 $x \in \mathbb{C}^n$ 成立。

由于 $A A^* - I$ 是 Hermite 矩阵, 故由

$$x^* (A A^* - I) x = 0, \quad \forall x$$

可得

$$A A^* = I.$$

于是 A 可逆, 并且

$$A^* = A^{-1}.$$

从而

$$A^* A = I.$$

注意到矩阵 $A^* A$ 的第 (i, j) 个元素为

$$a_i^* a_j = (\alpha_i, \alpha_j).$$

因此

$$(\alpha_i, \alpha_j) = \delta_{ij}.$$

这说明 $\alpha_1, \alpha_2, \dots, \alpha_n$ 两两正交, 且

$$\|\alpha_i\| = 1, \quad i = 1, 2, \dots, n.$$

所以 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是一组标准正交向量。

又因为 V 是 n 维空间, 而这组标准正交向量共有 n 个, 所以它们构成 V 的一组标准正交基。故 (i) 成立。

综上, 四个命题等价。 □

例题 1.5

- (1) 设 V 是酉空间, $\alpha, \beta \in V$ 。证明: α 与 β 正交的充分必要条件是, 对任意 $a, b \in \mathbb{C}$, 有

$$\|a\alpha + b\beta\|^2 = \|a\alpha\|^2 + \|b\beta\|^2.$$

- (2) 取酉空间 $V = \mathbb{C}^2$, 举例说明, 存在向量 $\alpha, \beta \in \mathbb{C}^2$, 使得 α 和 β 不正交, 但是

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2.$$

解答. (1) 先设 $\alpha \perp \beta$, 即

$$(\alpha, \beta) = 0.$$

则对任意 $a, b \in \mathbb{C}$,

$$\begin{aligned} \|a\alpha + b\beta\|^2 &= (a\alpha + b\beta, a\alpha + b\beta) \\ &= |a|^2\|\alpha\|^2 + |b|^2\|\beta\|^2 + a\bar{b}(\alpha, \beta) + b\bar{a}(\beta, \alpha) \\ &= |a|^2\|\alpha\|^2 + |b|^2\|\beta\|^2 \\ &= \|a\alpha\|^2 + \|b\beta\|^2. \end{aligned}$$

反过来, 若对任意 $a, b \in \mathbb{C}$, 都有

$$\|a\alpha + b\beta\|^2 = \|a\alpha\|^2 + \|b\beta\|^2,$$

则展开得

$$a\bar{b}(\alpha, \beta) + b\bar{a}(\beta, \alpha) = 0.$$

取 $a = b = 1$, 得到

$$(\alpha, \beta) + (\beta, \alpha) = 0,$$

即

$$2\operatorname{Re}(\alpha, \beta) = 0.$$

取 $a = 1, b = i$, 得到

$$\bar{i}(\alpha, \beta) + i(\beta, \alpha) = 0,$$

即

$$-i(\alpha, \beta) + i\overline{(\alpha, \beta)} = 0.$$

这推出

$$\operatorname{Im}(\alpha, \beta) = 0.$$

因此

$$(\alpha, \beta) = 0.$$

所以 $\alpha \perp \beta$ 。

(2) 在 $\mathbb{C}^2$ 中取

$$\alpha = (1, 0), \quad \beta = (i, 0).$$

则

$$(\alpha, \beta) = 1 \cdot \bar{i} + 0 \cdot \bar{0} = -i \neq 0,$$

因此 α 与 β 不正交。

但是

$$\alpha + \beta = (1 + i, 0),$$

所以

$$\|\alpha + \beta\|^2 = |1 + i|^2 = 2.$$

又

$$\|\alpha\|^2 = 1, \quad \|\beta\|^2 = 1,$$

故

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2.$$

□

例题 1.6 设 $V = \mathbb{R}[x]$, 对任意 $f(x), g(x) \in V$, 定义

$$(f(x), g(x)) = \int_0^1 f(x)g(x) dx,$$

则 V 成为欧氏空间。记

$$G = G(1, x, x^2, \dots, x^{n-1})$$

是 $1, x, x^2, \dots, x^{n-1}$ 的 Gram 矩阵, 证明:

$$G = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots & \frac{1}{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1} \end{pmatrix}.$$

该矩阵称为 Hilbert 矩阵。并证明

$$\det(G) > 0.$$

于是 Hilbert 矩阵是正定矩阵。

解答. 由 Gram 矩阵的定义,

$$G_{ij} = (x^{i-1}, x^{j-1}) = \int_0^1 x^{i-1}x^{j-1} dx = \int_0^1 x^{i+j-2} dx = \frac{1}{i+j-1}.$$

因此

$$G = \left(\frac{1}{i+j-1} \right)_{1 \leq i, j \leq n},$$

即题中所给矩阵。

下面证明 $\det(G) > 0$ 。对任意

$$c = (c_1, c_2, \dots, c_n)^T \in \mathbb{R}^n,$$

有

$$c^T G c = \sum_{i,j=1}^n c_i c_j (x^{i-1}, x^{j-1}) = \left(\sum_{i=1}^n c_i x^{i-1}, \sum_{j=1}^n c_j x^{j-1} \right).$$

记

$$p(x) = \sum_{i=1}^n c_i x^{i-1},$$

则

$$c^T G c = \int_0^1 p(x)^2 dx \geq 0.$$

若 $c^T G c = 0$ ，则

$$\int_0^1 p(x)^2 dx = 0.$$

由于 $p(x)^2$ 连续且非负，所以 $p(x) \equiv 0$ 于 $[0, 1]$ 。而多项式若在区间上恒为零，则它是零多项式，于是

$$c_1 = c_2 = \dots = c_n = 0.$$

因此对任意非零 $c \in \mathbb{R}^n$ ，有

$$c^T G c > 0.$$

故 G 是正定矩阵，从而

$$\det(G) > 0.$$

□

例题 1.7 在 $V = \mathbb{R}_n[x]$ 中定义内积

$$(f(x), g(x)) = \int_0^1 f(t)g(t) dt.$$

设

$$g_n(x) = x^n(1-x)^n,$$

再定义

$$f_n(x) = \frac{d^n}{dx^n} g_n(x),$$

则 $f_n(x)$ 是 n 次多项式。

(1) 证明：

(a)

$$\int_0^1 g_n(x) dx = \frac{(n!)^2}{(2n+1)!};$$

(b) 对 $i = 0, 1, \dots, n-1$, 有

$$(x^i, f_n(x)) = 0.$$

进而对任意 $f(x) \in \mathbb{R}_{n-1}[x]$, 有

$$(f(x), f_n(x)) = 0;$$

(c)

$$(f_n(x), f_n(x)) = \frac{(n!)^2}{2n+1}.$$

(2) 在 $\mathbb{R}_n[x]$ 中求与多项式 $1, x, \dots, x^{n-1}$ 都正交的单位向量;

(3) 求 $\mathbb{R}_n[x]$ 的一组标准正交基。

解答. 首先注意

$$g_n(x) = x^n(1-x)^n$$

是 $2n$ 次多项式, 其首项为

$$(-1)^n x^{2n}.$$

因此

$$f_n(x) = g_n^{(n)}(x)$$

是 n 次多项式, 且其首项系数为

$$(-1)^n \frac{(2n)!}{n!}.$$

(1) (a) 由 Beta 积分公式,

$$\int_0^1 x^n(1-x)^n dx = \frac{n!n!}{(2n+1)!}.$$

即

$$\int_0^1 g_n(x) dx = \frac{(n!)^2}{(2n+1)!}.$$

(b) 对 $i = 0, 1, \dots, n-1$, 有

$$(x^i, f_n(x)) = \int_0^1 x^i g_n^{(n)}(x) dx.$$

对右端分部积分。由于

$$g_n(x) = x^n(1-x)^n$$

在 $x = 0, 1$ 处有至少 n 阶零点, 所以 $g_n, g_n', \dots, g_n^{(n-1)}$ 在端点 $0, 1$ 处均为 0。因此分部积分时边界项全部为 0。

分部积分 $i+1$ 次后, x^i 的导数已经为 0, 于是

$$\int_0^1 x^i g_n^{(n)}(x) dx = 0.$$

故

$$(x^i, f_n(x)) = 0, \quad i = 0, 1, \dots, n-1.$$

任意 $f(x) \in \mathbb{R}_{n-1}[x]$ 都可以表示为

$$f(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1},$$

由内积的线性性,

$$(f(x), f_n(x)) = \sum_{i=0}^{n-1} a_i (x^i, f_n(x)) = 0.$$

(c) 由

$$f_n(x) = g_n^{(n)}(x),$$

有

$$(f_n, f_n) = \int_0^1 g_n^{(n)}(x) f_n(x) dx.$$

对右端分部积分 n 次。由于 $g_n, g'_n, \dots, g_n^{(n-1)}$ 在端点处均为 0, 边界项全部消失, 因此

$$\int_0^1 g_n^{(n)}(x) f_n(x) dx = (-1)^n \int_0^1 g_n(x) f_n^{(n)}(x) dx.$$

又

$$f_n^{(n)}(x) = g_n^{(2n)}(x).$$

因为 $g_n(x)$ 的首项为 $(-1)^n x^{2n}$, 所以

$$g_n^{(2n)}(x) = (-1)^n (2n)!.$$

从而

$$f_n^{(n)}(x) = (-1)^n (2n)!.$$

代入得

$$(f_n, f_n) = (-1)^n \int_0^1 g_n(x) \cdot (-1)^n (2n)! dx = (2n)! \int_0^1 g_n(x) dx.$$

由第 (a) 问,

$$(f_n, f_n) = (2n)! \cdot \frac{(n!)^2}{(2n+1)!} = \frac{(n!)^2}{2n+1}.$$

(2) 由第 (1) 问可知, $f_n(x)$ 与

$$1, x, \dots, x^{n-1}$$

都正交。又

$$\|f_n\|^2 = \frac{(n!)^2}{2n+1},$$

所以

$$\|f_n\| = \frac{n!}{\sqrt{2n+1}}.$$

因此所求单位向量可以取为

$$e_n(x) = \frac{\sqrt{2n+1}}{n!} f_n(x) = \frac{\sqrt{2n+1}}{n!} \frac{d^n}{dx^n} [x^n(1-x)^n].$$

当然, 其相反数也满足要求。

(3) 对 $k = 0, 1, \dots, n$, 令

$$g_k(x) = x^k(1-x)^k, \quad f_k(x) = \frac{d^k}{dx^k} g_k(x),$$

并定义

$$e_k(x) = \frac{\sqrt{2k+1}}{k!} f_k(x) = \frac{\sqrt{2k+1}}{k!} \frac{d^k}{dx^k} [x^k(1-x)^k].$$

当 $k = 0$ 时,

$$e_0(x) = 1.$$

由第 (1) 问, 对每个 k , $f_k(x)$ 与所有次数小于 k 的多项式正交。因此若 $j < k$, 则

$$f_j(x) \in \mathbb{R}_j[x] \subseteq \mathbb{R}_{k-1}[x],$$

从而

$$(f_j, f_k) = 0.$$

所以

$$e_0(x), e_1(x), \dots, e_n(x)$$

两两正交。

又由第 (1) 问,

$$\|f_k\|^2 = \frac{(k!)^2}{2k+1},$$

因此

$$\|e_k\| = 1.$$

于是

$$e_k(x) = \frac{\sqrt{2k+1}}{k!} \frac{d^k}{dx^k} [x^k(1-x)^k], \quad k = 0, 1, \dots, n$$

是 $\mathbb{R}_n[x]$ 的一组标准正交基。

□

2 Gram 矩阵、正规矩阵与酉对角化

例题 2.1 设矩阵 $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times p}$ 。

(1) 证明:

$$\|AB\| \leq \|A\| \|B\|,$$

这里矩阵的长度分别取的是对应的矩阵空间的标准内积；

(2) 如果 $m = n$ ，且 $\lambda \in \mathbb{C}$ 是 A 的一个特征值，证明：

$$|\lambda| \leq \|A\|.$$

解答. 这里矩阵空间取标准内积

$$(A, B) = \text{tr}(AB^*),$$

对应的范数为 Frobenius 范数：

$$\|A\|^2 = \sum_{i,j} |a_{ij}|^2.$$

(1) 先证明对任意 $x \in \mathbb{C}^n$ ，有

$$\|Ax\| \leq \|A\|\|x\|.$$

若 A 的第 i 行为 r_i ，则

$$(Ax)_i = \sum_{j=1}^n a_{ij}x_j.$$

由 Cauchy–Schwarz 不等式，

$$|(Ax)_i|^2 \leq \left(\sum_{j=1}^n |a_{ij}|^2 \right) \left(\sum_{j=1}^n |x_j|^2 \right).$$

对 i 求和，得到

$$\|Ax\|^2 = \sum_{i=1}^m |(Ax)_i|^2 \leq \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right) \|x\|^2 = \|A\|^2 \|x\|^2.$$

因此

$$\|Ax\| \leq \|A\|\|x\|.$$

设 B 的列向量为

$$B = (b_1, b_2, \dots, b_p),$$

则

$$AB = (Ab_1, Ab_2, \dots, Ab_p).$$

于是

$$\begin{aligned} \|AB\|^2 &= \sum_{j=1}^p \|Ab_j\|^2 \\ &\leq \sum_{j=1}^p \|A\|^2 \|b_j\|^2 \\ &= \|A\|^2 \sum_{j=1}^p \|b_j\|^2 \\ &= \|A\|^2 \|B\|^2. \end{aligned}$$

故

$$\|AB\| \leq \|A\|\|B\|.$$

(2) 若 λ 是 A 的特征值, 则存在非零向量 $x \in \mathbb{C}^n$, 使得

$$Ax = \lambda x.$$

由第 (1) 问中的结论,

$$\|\lambda x\| = \|Ax\| \leq \|A\|\|x\|.$$

即

$$|\lambda|\|x\| \leq \|A\|\|x\|.$$

因为 $x \neq 0$, 所以 $\|x\| > 0$, 从而

$$|\lambda| \leq \|A\|.$$

□

例题 2.2 设 $\xi_1, \xi_2, \dots, \xi_n$ 是 n 维内积空间 V 的一组标准正交基, $\alpha_1, \alpha_2, \dots, \alpha_n \in V$. 设 α_j 在这组标准正交基下的坐标为

$$X_j = (x_{1j}, x_{2j}, \dots, x_{nj})^T, \quad j = 1, 2, \dots, n.$$

证明:

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_n) = |\det(x_{ij})|^2.$$

解答. 记

$$X = (x_{ij})_{n \times n},$$

则 X 的第 j 列正是 α_j 在标准正交基 $\xi_1, \xi_2, \dots, \xi_n$ 下的坐标。于是

$$\alpha_j = \sum_{i=1}^n x_{ij} \xi_i.$$

因此 Gram 矩阵的第 (j, k) 个元素为

$$\begin{aligned} G_{jk} &= (\alpha_j, \alpha_k) \\ &= \left(\sum_{i=1}^n x_{ij} \xi_i, \sum_{\ell=1}^n x_{\ell k} \xi_\ell \right) \\ &= \sum_{i, \ell=1}^n x_{ij} \overline{x_{\ell k}} (\xi_i, \xi_\ell). \end{aligned}$$

由于 $\xi_1, \xi_2, \dots, \xi_n$ 是标准正交基,

$$(\xi_i, \xi_\ell) = \delta_{i\ell},$$

所以

$$G_{jk} = \sum_{i=1}^n x_{ij} \overline{x_{ik}}.$$

这说明

$$G = X^T \overline{X}.$$

于是

$$\det G = \det(X^T \overline{X}) = \det(X^T) \det(\overline{X}) = \det X \cdot \overline{\det X} = |\det X|^2.$$

即

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_n) = |\det(x_{ij})|^2.$$

□

例题 2.3 设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 n 维内积空间 V 的一组基。对其施行 Gram-Schmidt 正交化, 得到 V 的正交基

$$\beta_1, \beta_2, \dots, \beta_n.$$

证明:

$$\|\beta_j\|^2 = \frac{\det G(\alpha_1, \alpha_2, \dots, \alpha_j)}{\det G(\alpha_1, \alpha_2, \dots, \alpha_{j-1})}, \quad j = 1, 2, \dots, n,$$

其中约定零个向量的 Gram 方阵的行列式为 1。

解答. 由 Gram-Schmidt 正交化过程可知, 对每个 j , 有

$$\text{span}(\alpha_1, \alpha_2, \dots, \alpha_j) = \text{span}(\beta_1, \beta_2, \dots, \beta_j),$$

并且

$$\alpha_j = \beta_j + \sum_{i=1}^{j-1} c_{ij} \beta_i.$$

因此存在一个上三角矩阵 T_j , 其对角线元素全为 1, 使得

$$(\alpha_1, \alpha_2, \dots, \alpha_j) = (\beta_1, \beta_2, \dots, \beta_j) T_j.$$

于是 Gram 矩阵满足

$$G(\alpha_1, \alpha_2, \dots, \alpha_j) = T_j^T G(\beta_1, \beta_2, \dots, \beta_j) \overline{T_j}.$$

因为 T_j 是对角线全为 1 的上三角矩阵, 所以

$$\det T_j = 1.$$

故

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_j) = \det G(\beta_1, \beta_2, \dots, \beta_j).$$

又因为 $\beta_1, \beta_2, \dots, \beta_j$ 两两正交, 所以它们的 Gram 矩阵为对角矩阵:

$$G(\beta_1, \beta_2, \dots, \beta_j) = \text{diag}(\|\beta_1\|^2, \|\beta_2\|^2, \dots, \|\beta_j\|^2).$$

因此

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_j) = \prod_{i=1}^j \|\beta_i\|^2.$$

同理,

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_{j-1}) = \prod_{i=1}^{j-1} \|\beta_i\|^2.$$

两式相除, 得到

$$\|\beta_j\|^2 = \frac{\det G(\alpha_1, \alpha_2, \dots, \alpha_j)}{\det G(\alpha_1, \alpha_2, \dots, \alpha_{j-1})}.$$

当 $j = 1$ 时, 分母按约定为 1, 公式同样成立。 □

例题 2.4 设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 n 维内积空间 V 的一组向量, 证明:

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_n) \leq \|\alpha_1\|^2 \|\alpha_2\|^2 \cdots \|\alpha_n\|^2.$$

等号当且仅当 $\alpha_1, \alpha_2, \dots, \alpha_n$ 两两正交或其中含有零向量时成立。

由此证明: 对任意 $A = (a_{ij}) \in \mathbb{F}^{n \times n}$, 成立 Hadamard 不等式

$$|\det A|^2 \leq \prod_{j=1}^n \left(\sum_{i=1}^n |a_{ij}|^2 \right).$$

解答. 若 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性相关, 则

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_n) = 0,$$

而右边非负, 因此不等式成立。

下面设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性无关。对其施行 Gram-Schmidt 正交化, 得到正交基

$$\beta_1, \beta_2, \dots, \beta_n.$$

由上一题的证明可知

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_n) = \prod_{j=1}^n \|\beta_j\|^2.$$

又由 Gram-Schmidt 正交化的构造,

$$\beta_j = \alpha_j - \text{proj}_{\text{span}(\alpha_1, \dots, \alpha_{j-1})} \alpha_j.$$

因此 β_j 是 α_j 减去其在前面子空间上的正交投影所得向量, 于是

$$\|\beta_j\| \leq \|\alpha_j\|.$$

所以

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_n) = \prod_{j=1}^n \|\beta_j\|^2 \leq \prod_{j=1}^n \|\alpha_j\|^2.$$

接下来讨论等号。若某个 $\alpha_j = 0$, 则 Gram 矩阵有零行零列, 从而

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_n) = 0,$$

右边也为 0，故等号成立。

若所有 α_j 均非零，则等号成立当且仅当

$$\|\beta_j\| = \|\alpha_j\|, \quad j = 1, 2, \dots, n.$$

这等价于 α_j 在

$$\text{span}(\alpha_1, \dots, \alpha_{j-1})$$

上的正交投影为零，即

$$\alpha_j \perp \text{span}(\alpha_1, \dots, \alpha_{j-1}).$$

因此等价于

$$\alpha_i \perp \alpha_j, \quad i \neq j.$$

故等号当且仅当这些向量两两正交，或者其中含有零向量时成立。

现在取矩阵 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ ，把 A 的第 j 列看作 $\mathbb{R}^n$ 中的向量

$$\alpha_j = (a_{1j}, a_{2j}, \dots, a_{nj})^T.$$

由前一题可知

$$\det G(\alpha_1, \alpha_2, \dots, \alpha_n) = |\det A|^2.$$

另一方面，

$$\|\alpha_j\|^2 = \sum_{i=1}^n |a_{ij}|^2.$$

代入 Gram 行列式不等式，得到

$$|\det A|^2 \leq \prod_{j=1}^n \left(\sum_{i=1}^n |a_{ij}|^2 \right).$$

这就是 Hadamard 不等式。 □

例题 2.5 设 $V = \mathbb{R}[x]$ ，则 V 在内积

$$(f(x), g(x)) = \int_0^1 f(x)g(x) dx$$

下成为 Euclid 空间。令

$$W = \text{span}\{1, x^2, x^4, x^6, \dots\}.$$

求 $W^\perp$ 。子空间 W 有正交补吗？

解答. 先求 $W^\perp$ 。设 $f(x) \in W^\perp$ ，记

$$f(x) = a_0 + a_1x + \dots + a_mx^m.$$

由于 $f(x)$ 与 W 中任意向量正交，特别地，对 $k = 0, 1, \dots, m$ ，有

$$(f(x), x^{2k}) = 0.$$

也就是

$$\int_0^1 f(x)x^{2k} dx = 0, \quad k = 0, 1, \dots, m.$$

代入 $f(x)$ 得

$$\sum_{j=0}^m a_j \int_0^1 x^{j+2k} dx = 0,$$

即

$$\sum_{j=0}^m \frac{a_j}{j+2k+1} = 0, \quad k = 0, 1, \dots, m.$$

这给出了关于 $a_0, a_1, \dots, a_m$ 的线性方程组, 其系数矩阵为

$$\left(\frac{1}{j+2k+1} \right)_{0 \leq k, j \leq m}.$$

这是一个 Cauchy 型矩阵。因为

$$j+2k+1 = (j+1) + 2k,$$

其中 $j+1$ 两两不同, $2k$ 也两两不同, 所以该 Cauchy 矩阵行列式非零。因此线性方程组只有零解:

$$a_0 = a_1 = \dots = a_m = 0.$$

于是

$$f(x) = 0.$$

故

$$W^\perp = \{0\}.$$

但是 W 并不等于 V 。例如

$$x \notin W,$$

因为 W 中的元素都是偶次幂的线性组合, 即只含偶次幂的多项式。

若 W 有正交补, 则应有

$$V = W \oplus W^\perp.$$

但现在

$$W^\perp = \{0\},$$

于是右边为

$$W \oplus W^\perp = W,$$

这不可能等于 V 。

因此

$$W^\perp = \{0\}, \quad W \text{ 没有正交补.}$$

□

知识 2.1: 正规矩阵

设 $A \in M_n(\mathbb{C})$ 。若

$$A^*A = AA^*,$$

则称 A 为正规矩阵, 其中 A^* 表示 A 的共轭转置。

正规矩阵是一类能够被酉矩阵对角化的矩阵。其核心性质可以概括为:

$$A \text{ 正规} \iff A \text{ 可酉对角化}.$$

也就是说, 存在酉矩阵 U , 使得

$$U^*AU = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的特征值。

性质	说明
酉对角化	A 正规当且仅当 A 存在一组标准正交的特征向量。换言之, A 可以在某组标准正交基下表示为对角矩阵。
特征向量正交性	若 A 正规, 且 $\lambda \neq \mu$, 则属于不同特征值的特征子空间相互正交: $E_\lambda \perp E_\mu.$
范数相等	若 A 正规, 则对任意 $x \in \mathbb{C}^n$, 有 $\ Ax\ = \ A^*x\ .$
常见例子	Hermite 矩阵、酉矩阵、实对称矩阵、斜 Hermite 矩阵都是正规矩阵。

常见的正规矩阵包括:

$$A = A^* \implies A \text{ 是 Hermite 矩阵},$$

$$A^*A = I \implies A \text{ 是酉矩阵},$$

$$A^* = -A \implies A \text{ 是斜 Hermite 矩阵}.$$

这些矩阵都满足

$$A^*A = AA^*.$$

因此, 正规矩阵可以看作复内积空间中最适合用正交方法处理的一类矩阵。它们的本质特点是: 不仅可以对角化, 而且可以通过保持内积的酉变换对角化。

例题 2.6 证明:

- (1) n 维 Euclid 空间 V 上的线性变换 $\mathcal{A}$ 是正规变换的充分必要条件是, 存在多项式 $f(\lambda) \in \mathbb{R}[\lambda]$, 使得

$$\mathcal{A}^* = f(\mathcal{A});$$

- (2) n 阶实方阵 A 为正规方阵的充分必要条件是, 存在多项式 $f(\lambda) \in \mathbb{R}[\lambda]$, 使得

$$A^T = f(A).$$

解答. (1) 先证明必要性. 设 $\mathcal{A}$ 是 n 维 Euclid 空间 V 上的正规变换, 即

$$\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}.$$

由实正规变换的标准形可知, 存在 V 的一组标准正交基, 使得 $\mathcal{A}$ 的矩阵为分块对角形

$$\text{diag}(\lambda_1, \dots, \lambda_r, B_1, \dots, B_s),$$

其中 $\lambda_i \in \mathbb{R}$, 而每个 B_j 为形如

$$B_j = \begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix}, \quad b_j \neq 0$$

的 2 阶实方阵。

在该标准正交基下,

$$\mathcal{A}^*$$

的矩阵就是上述矩阵的转置, 即

$$\text{diag}(\lambda_1, \dots, \lambda_r, B_1^T, \dots, B_s^T),$$

其中

$$B_j^T = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}.$$

从复数角度看, B_j 对应复数

$$a_j + b_j i,$$

而 B_j^T 对应其共轭复数

$$a_j - b_j i.$$

因此只需找到一个实系数多项式 $f(\lambda)$, 使它在 $\mathcal{A}$ 的每个实特征值 λ_i 上取值为 λ_i , 并且在每个复特征值 $a_j + b_j i$ 上取值为 $a_j - b_j i$ 。

由于 $\mathcal{A}$ 是正规变换, 其最小多项式无重因子, 且在 $\mathbb{R}$ 上可以分解为若干互异的一次因子和不可约二次因子。由中国剩余定理, 存在 $f(\lambda) \in \mathbb{R}[\lambda]$, 使得

$$f(\lambda) \equiv \lambda$$

在每个实一次因子模下成立, 并且

$$f(\lambda) \equiv 2a_j - \lambda$$

在每个二次因子

$$(\lambda - a_j)^2 + b_j^2$$

模下成立。

对于二次块 B_j , 因为

$$f(\lambda) \equiv 2a_j - \lambda \pmod{(\lambda - a_j)^2 + b_j^2},$$

所以

$$f(B_j) = 2a_j I - B_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} = B_j^T.$$

对于实特征值块显然有

$$f(\lambda_i) = \lambda_i.$$

因此整体上得到

$$\mathcal{A}^* = f(\mathcal{A}).$$

再证明充分性。若存在 $f(\lambda) \in \mathbb{R}[\lambda]$, 使得

$$\mathcal{A}^* = f(\mathcal{A}),$$

则

$$\mathcal{A}\mathcal{A}^* = \mathcal{A}f(\mathcal{A}) = f(\mathcal{A})\mathcal{A} = \mathcal{A}^*\mathcal{A},$$

因为 $\mathcal{A}$ 与自身的任意多项式可交换。故 $\mathcal{A}$ 是正规变换。

综上,

$$\mathcal{A} \text{ 正规} \iff \exists f(\lambda) \in \mathbb{R}[\lambda], \quad \mathcal{A}^* = f(\mathcal{A}).$$

(2) 在 $\mathbb{R}^n$ 中取标准内积, 则实方阵 A 所对应线性变换的伴随变换矩阵为

$$A^T.$$

因此 A 为正规方阵, 即

$$AA^T = A^T A,$$

当且仅当对应线性变换 $\mathcal{A}$ 为正规变换。

由第 (1) 问,

$$\mathcal{A} \text{ 正规} \iff \exists f(\lambda) \in \mathbb{R}[\lambda], \quad \mathcal{A}^* = f(\mathcal{A}).$$

翻译成矩阵形式, 就是

$$A^T = f(A).$$

因此

$$A \text{ 为正规方阵} \iff \exists f(\lambda) \in \mathbb{R}[\lambda], \quad A^T = f(A).$$

□

例题 2.7 证明：两两可交换的 n 阶正规方阵集合可以同时酉相似于对角形。

解答. 设这族正规方阵为

$$\mathcal{S} = \{A_\lambda \mid \lambda \in \Lambda\},$$

且任意两个矩阵可交换，即

$$A_\lambda A_\mu = A_\mu A_\lambda, \quad \lambda, \mu \in \Lambda.$$

我们证明存在一个酉矩阵 U ，使得对任意 $\lambda \in \Lambda$ ，矩阵

$$U^* A_\lambda U$$

都是对角矩阵。

对维数 n 作归纳。

当 $n = 1$ 时结论显然成立。

设 $n > 1$ 。若 $\mathcal{S}$ 中每个矩阵都是数量矩阵，则它们在任意标准正交基下都已经是对角矩阵，结论成立。

否则，存在某个 $A \in \mathcal{S}$ 不是数量矩阵。由于 A 是正规矩阵，由正规矩阵的酉对角化定理，存在酉矩阵 U_1 ，使得

$$U_1^* A U_1 = \text{diag}(\lambda_1 I_{r_1}, \lambda_2 I_{r_2}, \dots, \lambda_s I_{r_s}),$$

其中

$$\lambda_1, \lambda_2, \dots, \lambda_s$$

两两不同，且 $s \geq 2$ 。

为了简化记号，经过一次酉相似后，不妨直接设

$$A = \text{diag}(\lambda_1 I_{r_1}, \lambda_2 I_{r_2}, \dots, \lambda_s I_{r_s}).$$

记对应的特征子空间为

$$V_i = \ker(A - \lambda_i I), \quad i = 1, 2, \dots, s.$$

则

$$\mathbb{C}^n = V_1 \oplus V_2 \oplus \dots \oplus V_s$$

是正交直和。

对任意 $B \in \mathcal{S}$ ，由于

$$AB = BA,$$

若 $v \in V_i$ ，即

$$Av = \lambda_i v,$$

则

$$A(Bv) = B(Av) = B(\lambda_i v) = \lambda_i Bv.$$

因此

$$Bv \in V_i.$$

这说明每个 V_i 都是所有 $B \in \mathcal{S}$ 的不变子空间。

于是对任意 $B \in \mathcal{S}$ ，它在分解

$$\mathbb{C}^n = V_1 \oplus V_2 \oplus \cdots \oplus V_s$$

下具有分块对角形式：

$$B = \text{diag}(B_1, B_2, \cdots, B_s).$$

又因为 B 是正规矩阵，所以

$$BB^* = B^*B.$$

写成分块形式，得到每个分块 B_i 也都是正规矩阵。并且若 $B, C \in \mathcal{S}$ ，由

$$BC = CB$$

可知在每个 V_i 上的限制也满足

$$B_i C_i = C_i B_i.$$

也就是说，在每个维数较小的空间 V_i 上，限制矩阵族仍然是一族两两可交换的正规矩阵。

由归纳假设，在每个 V_i 中存在一组标准正交基，使得所有限制矩阵同时对角化。将这些标准正交基合并起来，得到 $\mathbb{C}^n$ 的一组标准正交基。在这组基下，每个 $B \in \mathcal{S}$ 都是对角矩阵。

因此存在酉矩阵 U ，使得对任意 $B \in \mathcal{S}$ ，

$$U^* B U$$

都是对角矩阵。

故两两可交换的 n 阶正规方阵集合可以同时酉相似于对角形。 $\square$

3 欧氏几何、正交变换与角图

例题 3.1 证明： n 维欧氏空间中存在且至多存在 $n+1$ 个向量，使得它们两两之间夹成钝角。

解答. 这里“两两之间夹成钝角”等价于：对任意 $i \neq j$ ，有

$$(\alpha_i, \alpha_j) < 0.$$

先证明存在性。在 $\mathbb{R}^{n+1}$ 中取标准基

$$e_1, e_2, \cdots, e_{n+1},$$

并令

$$u = \frac{1}{n+1}(e_1 + e_2 + \cdots + e_{n+1}).$$

定义

$$v_i = e_i - u, \quad i = 1, 2, \cdots, n+1.$$

这些向量都属于超平面

$$H = \{x = (x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1 + x_2 + \dots + x_{n+1} = 0\}.$$

而 H 是一个 n 维欧氏空间。

当 $i \neq j$ 时,

$$(v_i, v_j) = (e_i - u, e_j - u) = (e_i, e_j) - (e_i, u) - (u, e_j) + (u, u).$$

由于

$$(e_i, e_j) = 0, \quad (e_i, u) = \frac{1}{n+1}, \quad (u, e_j) = \frac{1}{n+1}, \quad (u, u) = \frac{1}{n+1},$$

所以

$$(v_i, v_j) = -\frac{1}{n+1} < 0.$$

因此

$$v_1, v_2, \dots, v_{n+1}$$

两两之间夹成钝角。

下面证明至多性。设

$$\alpha_1, \alpha_2, \dots, \alpha_m$$

是 n 维欧氏空间中的一组非零向量, 且任意 $i \neq j$ 都有

$$(\alpha_i, \alpha_j) < 0.$$

我们证明 $m \leq n+1$ 。

考虑它们之间的线性关系。若存在一个非平凡线性关系

$$c_1 \alpha_1 + c_2 \alpha_2 + \dots + c_m \alpha_m = 0$$

其中系数既有正数又有负数, 则记

$$P = \{i \mid c_i > 0\}, \quad N = \{i \mid c_i < 0\}.$$

于是

$$\sum_{i \in P} c_i \alpha_i = \sum_{j \in N} (-c_j) \alpha_j.$$

记

$$w = \sum_{i \in P} c_i \alpha_i.$$

则

$$\begin{aligned} \|w\|^2 &= \left(\sum_{i \in P} c_i \alpha_i, \sum_{j \in N} (-c_j) \alpha_j \right) \\ &= \sum_{i \in P} \sum_{j \in N} c_i (-c_j) (\alpha_i, \alpha_j). \end{aligned}$$

由于 $c_i > 0$, $-c_j > 0$, 且 $(\alpha_i, \alpha_j) < 0$, 所以右端严格小于 0, 这与

$$\|w\|^2 \geq 0$$

矛盾。因此, 任意非平凡线性关系中的非零系数不可能同时有正有负。

于是所有非平凡线性关系的系数只能同号。这样的线性关系空间维数至多为 1; 否则若存在两个线性无关的关系, 可以作适当线性组合, 使得所得关系的系数既有正数又有负数, 矛盾。

因此向量组的线性关系空间维数至多为 1。也就是说

$$m - \text{rank}(\alpha_1, \dots, \alpha_m) \leq 1.$$

又

$$\text{rank}(\alpha_1, \dots, \alpha_m) \leq n,$$

所以

$$m \leq n + 1.$$

综上, n 维欧氏空间中存在且至多存在 $n + 1$ 个两两夹成钝角的向量。 $\square$

例题 3.2 设 V 是 n 维欧氏空间, 其中 $n \geq 2$ 。设 $\mathcal{A}$ 是 V 上的线性变换, 证明: 下面命题等价:

- (i) $\mathcal{A}$ 是 V 上的镜面反射, 即存在 V 中的超平面 W , 使得对任意 $\alpha = \beta + \gamma \in V$, 其中 $\beta \in W, \gamma \in W^\perp$, 有

$$\mathcal{A}(\alpha) = \beta - \gamma;$$

- (ii) 存在 V 的非零向量 η , 使得

$$\mathcal{A}(\alpha) = \alpha - 2 \frac{(\alpha, \eta)}{(\eta, \eta)} \eta, \quad \forall \alpha \in V;$$

- (iii) 存在 V 的标准正交基 $\eta_1, \eta_2, \dots, \eta_n$, 使得

$$\mathcal{A}(\eta_1) = -\eta_1, \quad \mathcal{A}(\eta_i) = \eta_i, \quad i = 2, \dots, n;$$

- (iv) $\mathcal{A}$ 在 V 的任意一个标准正交基下的矩阵有形式

$$E_n - 2\delta\delta^T,$$

其中 δ 为 $\mathbb{R}^n$ 的单位向量, 即 $\delta^T \delta = 1$ 。

解答. 证明四个命题循环等价。

- (i) $\Rightarrow$ (ii): 设 $\mathcal{A}$ 是关于超平面 W 的镜面反射。由于 W 是超平面, 所以

$$\dim W^\perp = 1.$$

取 $0 \neq \eta \in W^\perp$, 则

$$W = \{\alpha \in V \mid (\alpha, \eta) = 0\}.$$

对任意 $\alpha \in V$, 将其分解为

$$\alpha = \beta + \gamma, \quad \beta \in W, \quad \gamma \in W^\perp.$$

因为 $W^\perp = \text{span}\{\eta\}$, 所以

$$\gamma = c\eta.$$

两边与 η 作内积, 得

$$(\alpha, \eta) = (\beta, \eta) + c(\eta, \eta) = c(\eta, \eta).$$

因此

$$c = \frac{(\alpha, \eta)}{(\eta, \eta)},$$

从而

$$\gamma = \frac{(\alpha, \eta)}{(\eta, \eta)}\eta.$$

于是

$$\mathcal{A}(\alpha) = \beta - \gamma = (\alpha - \gamma) - \gamma = \alpha - 2\gamma = \alpha - 2\frac{(\alpha, \eta)}{(\eta, \eta)}\eta.$$

(ii) $\Rightarrow$ (iii): 设

$$\mathcal{A}(\alpha) = \alpha - 2\frac{(\alpha, \eta)}{(\eta, \eta)}\eta, \quad \eta \neq 0.$$

令

$$\eta_1 = \frac{\eta}{\|\eta\|}.$$

将 η_1 扩充为 V 的一组标准正交基

$$\eta_1, \eta_2, \dots, \eta_n.$$

则

$$\mathcal{A}(\eta_1) = \eta_1 - 2\frac{(\eta_1, \eta)}{(\eta, \eta)}\eta = \eta_1 - 2\eta_1 = -\eta_1.$$

而当 $i = 2, \dots, n$ 时, $\eta_i \perp \eta$, 所以

$$\mathcal{A}(\eta_i) = \eta_i - 2\frac{(\eta_i, \eta)}{(\eta, \eta)}\eta = \eta_i.$$

故 (iii) 成立。

(iii) $\Rightarrow$ (i): 设存在标准正交基 $\eta_1, \eta_2, \dots, \eta_n$, 使得

$$\mathcal{A}(\eta_1) = -\eta_1, \quad \mathcal{A}(\eta_i) = \eta_i, \quad i = 2, \dots, n.$$

令

$$W = \text{span}\{\eta_2, \dots, \eta_n\}.$$

则 W 是超平面, 且

$$W^\perp = \text{span}\{\eta_1\}.$$

任意 $\alpha \in V$ 可唯一写成

$$\alpha = \beta + \gamma, \quad \beta \in W, \quad \gamma \in W^\perp.$$

由 $\mathcal{A}$ 在上述基上的作用可知

$$\mathcal{A}(\beta) = \beta, \quad \mathcal{A}(\gamma) = -\gamma.$$

因此

$$\mathcal{A}(\alpha) = \beta - \gamma.$$

故 $\mathcal{A}$ 是关于超平面 W 的镜面反射。

(ii) $\Leftrightarrow$ (iv): 取 V 的任意一个标准正交基。设非零向量 η 在该基下的坐标列为 u , 令

$$\delta = \frac{u}{\|u\|}.$$

则 $\delta^T \delta = 1$, 且公式

$$\mathcal{A}(\alpha) = \alpha - 2 \frac{(\alpha, \eta)}{(\eta, \eta)} \eta$$

在坐标下正好写成

$$X \mapsto X - 2\delta(\delta^T X) = (E_n - 2\delta\delta^T)X.$$

因此 $\mathcal{A}$ 的矩阵为

$$E_n - 2\delta\delta^T.$$

反过来, 若 $\mathcal{A}$ 在某一标准正交基下的矩阵为

$$E_n - 2\delta\delta^T, \quad \delta^T \delta = 1,$$

取坐标为 δ 的向量 η , 则 $\|\eta\| = 1$, 并且对任意 $\alpha \in V$,

$$\mathcal{A}(\alpha) = \alpha - 2(\alpha, \eta)\eta = \alpha - 2 \frac{(\alpha, \eta)}{(\eta, \eta)} \eta.$$

故 (ii) 成立。

综上, 四个命题等价。 □

例题 3.3 设 $\mathcal{A}$ 是 n 维欧氏空间 V 上的正交变换, 其中 $n \geq 2$ 。设 1 是 $\mathcal{A}$ 的特征值, 且属于特征值 1 的特征子空间 V_1 的维数为 $n-1$ 。证明: $\mathcal{A}$ 是关于超平面 V_1 的镜面反射。

解答. 因为

$$\dim V_1 = n-1,$$

所以 V_1 是 V 中的一个超平面。又因为 $\mathcal{A}$ 是正交变换, 所以保持内积。

取单位向量 $\eta \in V_1^\perp$ 。由于 $V_1^\perp$ 是一维子空间, 只需考察 $\mathcal{A}(\eta)$ 。对任意 $\beta \in V_1$, 有

$$(\mathcal{A}\eta, \beta) = (\mathcal{A}\eta, \mathcal{A}\beta) = (\eta, \beta) = 0.$$

因此

$$\mathcal{A}\eta \in V_1^\perp.$$

又 $\mathcal{A}$ 保持长度, 所以

$$\|\mathcal{A}\eta\| = \|\eta\| = 1.$$

由于 $V_1^\perp = \text{span}\{\eta\}$, 于是

$$\mathcal{A}\eta = \eta \quad \text{或} \quad \mathcal{A}\eta = -\eta.$$

若 $\mathcal{A}\eta = \eta$, 则 η 也属于特征值 1 的特征子空间, 于是

$$\dim V_1 = n,$$

这与题设 $\dim V_1 = n - 1$ 矛盾。因此

$$\mathcal{A}\eta = -\eta.$$

于是 $\mathcal{A}$ 在 V_1 上恒等, 在 $V_1^\perp$ 上取相反数。也就是说, 任意

$$\alpha = \beta + \gamma, \quad \beta \in V_1, \quad \gamma \in V_1^\perp,$$

都有

$$\mathcal{A}(\alpha) = \beta - \gamma.$$

故 $\mathcal{A}$ 是关于超平面 V_1 的镜面反射。 □

例题 3.4 设 V 是 n 维欧氏空间, 其中 $n \geq 2$ 。证明: V 上的第二类正交变换是一个镜面反射, 或者是一个镜面反射和一个第一类正交变换的乘积。

解答. 设 $\mathcal{A}$ 是 V 上的第二类正交变换, 即

$$\det \mathcal{A} = -1.$$

若 $\mathcal{A}$ 本身就是镜面反射, 则结论成立。

否则, 任取一个镜面反射 $\mathcal{R}$ 。镜面反射在某组标准正交基下的矩阵为

$$\text{diag}(-1, 1, \dots, 1),$$

因此

$$\det \mathcal{R} = -1.$$

考虑

$$\mathcal{B} = \mathcal{R}\mathcal{A}.$$

由于正交变换的乘积仍是正交变换, $\mathcal{B}$ 是正交变换。并且

$$\det \mathcal{B} = \det \mathcal{R} \det \mathcal{A} = (-1)(-1) = 1.$$

所以 $\mathcal{B}$ 是第一类正交变换。

由

$$\mathcal{B} = \mathcal{R}\mathcal{A}$$

得

$$\mathcal{A} = \mathcal{R}^{-1}\mathcal{B}.$$

而镜面反射满足

$$\mathcal{R}^{-1} = \mathcal{R}.$$

因此

$$\mathcal{A} = \mathcal{R}\mathcal{B},$$

即 $\mathcal{A}$ 是一个镜面反射和一个第一类正交变换的乘积。

综上, 第二类正交变换要么本身是镜面反射, 要么可以表示为一个镜面反射和一个第一类正交变换的乘积。 $\square$

例题 3.5 设 V 是 n 维欧氏空间, 其中 $n \geq 2$ 。

(1) 设 $\alpha, \beta \in V$ 是两个不同的向量, 且

$$\|\alpha\| = \|\beta\|.$$

证明: 存在 V 的镜面反射 $\mathcal{A}$, 使得

$$\mathcal{A}(\alpha) = \beta.$$

(2) 证明: V 上的任一正交变换可以表示成至多 $n+1$ 个镜面反射的乘积。

解答. (1) 因为 $\alpha \neq \beta$, 令

$$\eta = \alpha - \beta.$$

则 $\eta \neq 0$. 考虑关于法向量 η 的镜面反射

$$\mathcal{A}(x) = x - 2\frac{(x, \eta)}{(\eta, \eta)}\eta.$$

我们证明 $\mathcal{A}(\alpha) = \beta$ 。

由于

$$\eta = \alpha - \beta,$$

有

$$(\alpha, \eta) = (\alpha, \alpha - \beta) = \|\alpha\|^2 - (\alpha, \beta).$$

又

$$(\eta, \eta) = \|\alpha - \beta\|^2 = \|\alpha\|^2 - 2(\alpha, \beta) + \|\beta\|^2.$$

因为 $\|\alpha\| = \|\beta\|$, 所以

$$(\eta, \eta) = 2(\|\alpha\|^2 - (\alpha, \beta)) = 2(\alpha, \eta).$$

因此

$$2\frac{(\alpha, \eta)}{(\eta, \eta)} = 1.$$

从而

$$\mathcal{A}(\alpha) = \alpha - \eta = \alpha - (\alpha - \beta) = \beta.$$

(2) 设 $\mathcal{T}$ 是 V 上的正交变换。取 V 的一组标准正交基

$$e_1, e_2, \dots, e_n.$$

我们用归纳构造镜面反射, 使得逐步把 $\mathcal{T}e_i$ 送回 e_i 。

首先, $\mathcal{T}$ 保持长度, 所以

$$\|\mathcal{T}e_1\| = \|e_1\| = 1.$$

若 $\mathcal{T}e_1 \neq e_1$, 由第 (1) 问, 存在镜面反射 $\mathcal{R}_1$, 使得

$$\mathcal{R}_1\mathcal{T}e_1 = e_1.$$

若 $\mathcal{T}e_1 = e_1$, 则令 $\mathcal{R}_1$ 为恒等变换, 不需要使用镜面反射。

设已经构造出若干镜面反射, 使得

$$\mathcal{S}_{k-1}\mathcal{T}e_i = e_i, \quad i = 1, 2, \dots, k-1.$$

其中 $\mathcal{S}_{k-1}$ 是至多 $k-1$ 个镜面反射的乘积。考虑

$$u = \mathcal{S}_{k-1}\mathcal{T}e_k.$$

因为 $\mathcal{S}_{k-1}\mathcal{T}$ 是正交变换, 所以

$$\|u\| = 1.$$

且对于 $i < k$,

$$(u, e_i) = (\mathcal{S}_{k-1}\mathcal{T}e_k, \mathcal{S}_{k-1}\mathcal{T}e_i) = (e_k, e_i) = 0.$$

因此 u 与 $e_1, \dots, e_{k-1}$ 都正交, 且 e_k 也与它们正交。于是 u, e_k 都属于子空间

$$U_k = \text{span}\{e_1, \dots, e_{k-1}\}^\perp.$$

若 $u \neq e_k$, 则在 U_k 中由第 (1) 问可取一个镜面反射, 把 u 变成 e_k ; 将它 $U_k^\perp$ 上定义为恒等, 便得到 V 上的一个镜面反射 $\mathcal{R}_k$, 且不改变已经固定的 $e_1, \dots, e_{k-1}$ 。于是

$$\mathcal{R}_k\mathcal{S}_{k-1}\mathcal{T}e_i = e_i, \quad i = 1, \dots, k.$$

若 $u = e_k$, 则无需加入新的镜面反射。

如此进行到 $k = n$, 得到一个由至多 n 个镜面反射组成的乘积 $\mathcal{S}$, 使得

$$\mathcal{S}\mathcal{T}e_i = e_i, \quad i = 1, \dots, n.$$

因此

$$\mathcal{S}\mathcal{T} = I,$$

从而

$$\mathcal{T} = \mathcal{S}^{-1}.$$

由于镜面反射的逆仍是其自身, $\mathcal{S}^{-1}$ 仍是至多 n 个镜面反射的乘积。

因此 $\mathcal{T}$ 可以表示成至多 n 个镜面反射的乘积, 特别地, 也可以表示成至多 $n+1$ 个镜面反射的乘积。

□

例题 3.6 证明: 如果 n 阶正交方阵 Q 的行列式为 1, 则 Q 可以写为有限个形如

$$Q_{jk} = E_n + (\cos \theta - 1)(E_{jj} + E_{kk}) + \sin \theta(E_{jk} - E_{kj}), \quad 1 \leq j < k \leq n,$$

的方阵的乘积; 如果 Q 的行列式为 -1 , 则还应添上方阵

$$\text{diag}(1, \dots, 1, -1).$$

解答. 矩阵

$$Q_{jk} = E_n + (\cos \theta - 1)(E_{jj} + E_{kk}) + \sin \theta(E_{jk} - E_{kj})$$

只在第 j, k 个坐标平面上起作用, 其对应的 2 阶块为

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix},$$

其余坐标保持不变。因此 Q_{jk} 是正交矩阵, 且

$$\det Q_{jk} = 1.$$

这类矩阵称为 Givens 旋转矩阵。

先设

$$Q \in O(n), \quad \det Q = 1.$$

我们证明 Q 可以写成有限个 Q_{jk} 的乘积。

对 Q 的第一列向量作平面旋转消元。因为第一列是单位向量, 可以通过依次在 $(1, n), (1, n-1), \dots, (1, 2)$ 坐标平面内作适当的 Givens 旋转, 将第一列化为

$$(1, 0, \dots, 0)^T.$$

也就是说, 存在有限个 Givens 旋转矩阵 $R_1, \dots, R_s$, 使得

$$R_s \cdots R_1 Q = \begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix}.$$

由于左端仍为正交矩阵, 第一列为 e_1 , 故第一行除第一个元素外也必须为零, 于是

$$R_s \cdots R_1 Q = \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix},$$

其中 Q_1 是一个 $(n-1)$ 阶正交矩阵。并且

$$\det Q_1 = \det(R_s \cdots R_1 Q) = 1.$$

这里用到每个 R_i 的行列式均为 1。

对 Q_1 重复同样过程，继续用 Givens 旋转将其化为

$$\begin{pmatrix} 1 & 0 \\ 0 & Q_2 \end{pmatrix},$$

其中 $Q_2 \in O(n-2)$ 且 $\det Q_2 = 1$ 。有限步后，得到

$$R_m \cdots R_1 Q = E_n,$$

其中每个 R_i 都是某个 Q_{jk} 型矩阵。于是

$$Q = R_1^{-1} \cdots R_m^{-1}.$$

而 $Q_{jk}(\theta)^{-1} = Q_{jk}(-\theta)$ ，仍是同一类型矩阵。因此 Q 可以写成有限个 Q_{jk} 型矩阵的乘积。

若

$$\det Q = -1,$$

令

$$D = \text{diag}(1, \cdots, 1, -1).$$

则

$$\det(QD) = \det Q \det D = (-1)(-1) = 1.$$

由上面已证结论， QD 可以写成有限个 Q_{jk} 型矩阵的乘积：

$$QD = R_1 R_2 \cdots R_m.$$

由于

$$D^{-1} = D,$$

故

$$Q = R_1 R_2 \cdots R_m D.$$

也就是说，当 $\det Q = -1$ 时，除了有限个平面旋转矩阵外，还需要添上

$$\text{diag}(1, \cdots, 1, -1).$$

综上，结论成立。 □

例题 3.7 设 V 是 Euclid 空间， S 是 V 的线性无关有限子集，且满足：对任意 $\alpha, \beta \in S, \alpha \neq \beta$ ，有

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} \leq 0,$$

且为整数。 S 的角图定义为：顶点为 S 中的所有向量；对任意 $\alpha, \beta \in S, \alpha \neq \beta$, α 和 β 之间连

$$\frac{4(\alpha, \beta)^2}{(\alpha, \alpha)(\beta, \beta)}$$

条边。

(1) 证明：对任意 $\alpha, \beta \in S, \alpha \neq \beta$, α 和 β 之间连的边数不超过 3，且当连的边数为 1 时， α 和 β 的长度相同；

(2) 如果 S 的角图为单链

$$\alpha_1 - \alpha_2 - \cdots - \alpha_n,$$

记

$$a = (\alpha_1, \alpha_1), \quad \alpha = \sum_{k=1}^n k \alpha_k,$$

证明：

$$(\alpha, \alpha) = \frac{n(n+1)}{2} a;$$

(3) 如果 S 的角图为

$$\alpha_1 - \alpha_2 - \cdots - \alpha_p \iff \beta_q - \beta_{q-1} - \cdots - \beta_1,$$

其中 $\iff$ 表示 α_p 与 β_q 之间有两条边，且 $p \geq q \geq 1$ ，证明：

$$p = q = 2, \quad \text{或者} \quad p \geq q = 1.$$

解答. (1) 对 $\alpha, \beta \in S, \alpha \neq \beta$ ，记

$$m = -\frac{2(\alpha, \beta)}{(\alpha, \alpha)}, \quad n = -\frac{2(\alpha, \beta)}{(\beta, \beta)}.$$

由题设可知 m, n 都是非负整数。于是 α 和 β 之间连的边数为

$$\frac{4(\alpha, \beta)^2}{(\alpha, \alpha)(\beta, \beta)} = mn.$$

由 Cauchy-Schwarz 不等式，

$$(\alpha, \beta)^2 \leq (\alpha, \alpha)(\beta, \beta).$$

由于 S 线性无关， α, β 不成比例，所以等号不能成立，因此

$$\frac{4(\alpha, \beta)^2}{(\alpha, \alpha)(\beta, \beta)} < 4.$$

而它又是非负整数，故只能为

$$0, 1, 2, 3.$$

因此连边数不超过 3。

若连边数为 1, 则

$$mn = 1.$$

因为 m, n 是非负整数, 所以

$$m = n = 1.$$

即

$$-\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = 1, \quad -\frac{2(\alpha, \beta)}{(\beta, \beta)} = 1.$$

从而

$$(\alpha, \alpha) = (\beta, \beta),$$

即 α 和 β 的长度相同。

- (2) 由于角图为单链, 任意相邻两个顶点之间只有一条边。由第 (1) 问可知, 相邻向量长度相同。因此

$$(\alpha_1, \alpha_1) = (\alpha_2, \alpha_2) = \cdots = (\alpha_n, \alpha_n) = a.$$

又因为相邻顶点之间连一条边, 且内积非正, 所以

$$(\alpha_k, \alpha_{k+1}) = -\frac{a}{2}, \quad k = 1, 2, \cdots, n-1.$$

非相邻顶点之间没有边, 因此

$$(\alpha_i, \alpha_j) = 0, \quad |i - j| \geq 2.$$

于是

$$\begin{aligned} (\alpha, \alpha) &= \left(\sum_{k=1}^n k\alpha_k, \sum_{k=1}^n k\alpha_k \right) \\ &= \sum_{k=1}^n k^2(\alpha_k, \alpha_k) + 2 \sum_{k=1}^{n-1} k(k+1)(\alpha_k, \alpha_{k+1}) \\ &= a \sum_{k=1}^n k^2 - a \sum_{k=1}^{n-1} k(k+1). \end{aligned}$$

计算得

$$\sum_{k=1}^n k^2 - \sum_{k=1}^{n-1} k(k+1) = n^2 - \sum_{k=1}^{n-1} k = n^2 - \frac{n(n-1)}{2} = \frac{n(n+1)}{2}.$$

因此

$$(\alpha, \alpha) = \frac{n(n+1)}{2}a.$$

- (3) 记

$$a = (\alpha_1, \alpha_1), \quad b = (\beta_1, \beta_1).$$

因为 $\alpha_1, \cdots, \alpha_p$ 构成单链, 所以由第 (2) 问, 令

$$u = \sum_{i=1}^p i\alpha_i,$$

则

$$(u, u) = \frac{p(p+1)}{2}a.$$

同理, 令

$$v = \sum_{j=1}^q j\beta_j,$$

则

$$(v, v) = \frac{q(q+1)}{2}b.$$

由于图中只有 α_p 与 β_q 相连, 且它们之间有两条边, 所以

$$\frac{4(\alpha_p, \beta_q)^2}{(\alpha_p, \alpha_p)(\beta_q, \beta_q)} = 2.$$

又 α_p 与其他 β_j 不相连, β_q 与其他 α_i 不相连, 因此

$$(u, v) = pq(\alpha_p, \beta_q).$$

于是

$$(u, v)^2 = p^2 q^2 (\alpha_p, \beta_q)^2 = p^2 q^2 \cdot \frac{ab}{2}.$$

由于 S 线性无关, 所以 u, v 线性无关。由 Cauchy-Schwarz 不等式的严格形式,

$$(u, v)^2 < (u, u)(v, v).$$

代入上式, 得到

$$p^2 q^2 \cdot \frac{ab}{2} < \frac{p(p+1)}{2}a \cdot \frac{q(q+1)}{2}b.$$

约去正数 $\frac{pqab}{4}$, 得

$$2pq < (p+1)(q+1).$$

即

$$pq - p - q < 1,$$

也就是

$$(p-1)(q-1) < 2.$$

又 $p \geq q \geq 1$ 。若 $q = 1$, 则得到

$$p \geq q = 1.$$

若 $q \geq 2$, 则

$$(p-1)(q-1) < 2$$

且 $p \geq q$, 只能有

$$p = q = 2.$$

因此

$$p = q = 2, \quad \text{或者} \quad p \geq q = 1.$$

□