

定义 7.2.8. 给定两个同阶方阵 A 和 B . 若存在可逆矩阵 P 使得 $A = P^T B P$, 则称 A 是 B 由相合变换 (congruent transformation) 得到的, 称 A 与 B 是相合的.

注 7.2.9. (1) 容易验证相合关系也是一个等价关系, 满足反身性, 对称性以及传递性.

(2) 上面的讨论说明: 内积在不同基下的矩阵是相合的.

(3) 上面的讨论也可以用来说明: 假定有两个矩阵 A 和 B 是相合的, 而 A 是内积在某组基下的度量矩阵, 那么 B 必然是同样的内积在另外一组基下的度量矩阵. 并且, 用于做相合变换的可逆矩阵构成了从第一组基到第二组基的过渡矩阵.

(4) 用正交矩阵来作相似变换与相合变换, 效果是一样的.

(5) 我们一般是在实方阵的框架下讨论相合. 不难看出, 定义中的 A 是实对称的, 当且仅当 B 是实对称的.

7.2.2 标准正交基

定义 7.2.10. 在欧氏空间 V 中, 由两两正交的非零向量构成的向量组称为一个正交向量组, 由正交向量组构成的基称为一组正交基, 由单位向量构成的正交基称为一组标准正交基, 单位正交基或规范正交基 (orthonormal basis).

例 7.2.11. (1) \mathbb{R}^n 的自然基在标准内积下是最常用的一组标准正交基.

(2) $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})^T, (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)^T, (\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}})^T$ 是 \mathbb{R}^3 的一组标准正交基.

(3) 在 $C[-\pi, \pi]$ 上定义内积

$$(f, g) := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x) dx.$$

对于任意给定的正整数 n , 不难验证集合

$$\left\{ \frac{1}{\sqrt{2}}, \cos(x), \cos(2x), \dots, \cos(nx), \sin(x), \sin(2x), \dots, \sin(nx) \right\}$$

是一个由单位向量构成的正交向量组. 由于已知 $C[-\pi, \pi]$ 是实数域上的无穷维线性空间, 故这组向量不可能构成 $C[-\pi, \pi]$ 的标准正交基.

命题 7.2.12. 若 $\alpha_1, \dots, \alpha_r$ 是欧氏空间的一个正交向量组, 那么它一定是线性无关的, 从而构成所生成的子空间的一组基.

证明. 由条件可知, 对于任意的 $i \neq j$, 有 $(\alpha_i, \alpha_j) = 0$. 若 $\sum_{i=1}^r \lambda_i \alpha_i = \mathbf{0}$, 则对于任意的 j , 有

$$0 = \left(\sum_{i=1}^r \lambda_i \alpha_i, \alpha_j \right) = \sum_{i=1}^r \lambda_i (\alpha_i, \alpha_j) = \lambda_j (\alpha_j, \alpha_j) = \lambda_j \|\alpha_j\|^2.$$

因为 α_j 不是零向量 (正交向量组的定义中的特别要求), $\|\alpha_j\| \neq 0$, 从而 $\lambda_j = 0$ 对于任意的 j 成立. 这说明该向量组是线性无关的. \square

推论 7.2.13. 在 n 维欧氏空间中两两正交的非零向量的个数不会超过 n 个.

回忆一下, 单位阵可以写成 $I_n = (\delta_{ij})_{n \times n}$, 其中 **Kronecker** 符号 δ_{ij} 是这样定义的:

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

推论 7.2.14. 向量组 $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ 构成 n 维欧氏空间 V 的标准正交基的充要条件是对于任意的 i, j , 有 $(\mathbf{x}_i, \mathbf{x}_j) = \delta_{ij}$.

定义 7.2.15. 对于欧氏空间 V 中的向量 $\mathbf{x}_1, \dots, \mathbf{x}_m$, 令 $\mathbf{G} = (g_{ij})_{m \times m}$, 其中 $g_{ij} = (\mathbf{x}_i, \mathbf{x}_j)$. 这样得到的对称矩阵 \mathbf{G} 称为向量组 $\mathbf{x}_1, \dots, \mathbf{x}_m$ 的 **Gram** 矩阵.

注 7.2.16. (1) 内积在 V 的基 $\alpha_1, \dots, \alpha_n$ 下的度量矩阵是该向量组的 *Gram* 矩阵.

- (2) 由定义不难看出, 向量组 $\mathbf{x}_1, \dots, \mathbf{x}_m$ 是一组由单位向量构成的正交组的充要条件是对于任意的 i, j , 有 $(\mathbf{x}_i, \mathbf{x}_j) = \delta_{ij}$. 用这儿的语言表述, 这等价于说, 它们的 *Gram* 矩阵为 m 阶的单位阵 I_m .
- (3) 若 V 是具有标准内积的欧氏空间 \mathbb{R}^n , 设 \mathbf{X} 是以 $\mathbf{x}_1, \dots, \mathbf{x}_m$ 为列向量的 $n \times m$ 矩阵, 那么定义中的 $\mathbf{G} = \mathbf{X}^\top \mathbf{X}$.
- (4) 在 (3) 中, 若 $m = n$ 时, 这意味着 $\mathbf{x}_1, \dots, \mathbf{x}_n$ 是数组空间 \mathbb{R}^n 的一组标准正交基的充要条件是以其为列向量组的矩阵 \mathbf{X} 满足 $\mathbf{X}^\top \mathbf{X} = I_n$, 即 \mathbf{X} 为正交矩阵. 进一步地, 由于 \mathbf{X} 是正交矩阵当且仅当 \mathbf{X}^\top 也是一个正交矩阵, 故这也等价于说 \mathbf{X} 的行向量组在数组空间 \mathbb{R}^n 的标准内积下构成一个正交向量组.
- (5) 仍然考虑 (3) 中的情形. 在例 5.5.7 中我们已经证明了 $\text{rank}(\mathbf{X}^\top \mathbf{X}) = \text{rank}(\mathbf{X})$. 由此可知, $\det(\mathbf{G}) \neq 0$ 当且仅当 \mathbf{G} 是满秩矩阵, 当且仅当 \mathbf{X} 是列满秩的, 而这也当且仅当 $\mathbf{x}_1, \dots, \mathbf{x}_m$ 线性无关.

(6) 一般而言, 欧氏空间 V 中的向量 $\mathbf{x}_1, \dots, \mathbf{x}_m$ 的 *Gram* 矩阵 \mathbf{G} 是半正定的对称矩阵, 并且 \mathbf{G} 是正定的当且仅当向量组 $\mathbf{x}_1, \dots, \mathbf{x}_m$ 线性无关 (参见教材最后一章最后一节的定义). 另外, 我们有 $\text{rank}(\mathbf{G}) = \text{rank}(\mathbf{x}_1, \dots, \mathbf{x}_m)$.

问题 7.2.17. 我们为什么要关心标准正交基呢?

注 7.2.18. 若 $\epsilon_1, \dots, \epsilon_n$ 是 V 的一组标准正交基, 向量 \mathbf{x} 在这组基下的坐标为 $(x_1, \dots, x_n)^T$, 即有 $\mathbf{x} = \sum_{i=1}^n x_i \epsilon_i$. 将该式与 ϵ_j 作内积, 我们得到

$$(\mathbf{x}, \epsilon_j) = \left(\sum_{i=1}^n x_i \epsilon_i, \epsilon_j \right) = \sum_{i=1}^n x_i (\epsilon_i, \epsilon_j) = x_j.$$

这说明向量坐标的分量可以用内积表达:

$$\mathbf{x} = \sum_{i=1}^n (\mathbf{x}, \epsilon_i) \epsilon_i.$$

该式称为 \mathbf{x} 的 *Fourier* 展开, 其中每个系数 (\mathbf{x}, ϵ_i) 都称为 \mathbf{x} 的 *Fourier* 系数. 设向量 $\mathbf{y} \in V$ 在这组标准正交基下的坐标为 $(y_1, \dots, y_n)^T$, 即有 $\mathbf{y} = \sum_{j=1}^n y_j \epsilon_j$. 此时,

$$(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^n x_i \epsilon_i, \sum_{j=1}^n y_j \epsilon_j \right) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j (\epsilon_i, \epsilon_j) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \delta_{ij} = \sum_{i=1}^n x_i y_i.$$

特别地, 我们有帕塞瓦尔等式

$$(\mathbf{x}, \mathbf{x}) = \sum_{i=1}^n x_i^2,$$

即向量的模满足勾股定理:

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

问题 7.2.19. 既然标准正交基有这么好的性质, 它是否一定存在呢?

注 7.2.20. 例 7.1.3 的 (2) 表明, 相对于有限维实线性空间 V 的任意一组基 $\alpha_1, \dots, \alpha_n$, V 上都存在一个相应的内积, 使得这组基在该内积下成为一组标准正交基.

显然, 上面的注并没有解答问题 7.2.19 中提出的疑问. 接下来我们说明, 有限维的欧氏空间在给定的内积下, 也必然存在一组标准正交基.

回忆. 有限维线性空间的任意一组线性无关的向量都可以扩充为向量空间的一组基.

类似地, 有限维欧氏空间的任意一个正交向量组都可以扩充为该欧氏空间的一组正交基. 为此, 我们打算证明如下的定理.

定理 7.2.21. 从 n 维欧氏空间 V 的任意一组基出发, 可以构造出一组标准正交基. 特别地, V 的标准正交基一定存在.

为了证明上面的结果, 我们需要先做一点准备.

讨论. (1) 若向量 $\alpha \in V$ 非零, 对其做归一化 (也称作标准化或单位化), 可以得到单位向量:

$$\alpha \mapsto \frac{\alpha}{\|\alpha\|}.$$

新得到的向量, 从几何意义上来说, 与原向量具有相同的方向.

(2) 在欧氏空间 V 中任意取定一个非零向量 α , 对于 V 的中任意向量 β , 我们希望找到

正交分解 $\beta = \beta_1 + \beta_2$, 其中 β_1 与 α 平行, 而 β_2 与 α 垂直.

为了找到这样的分解, 我们不妨设 $\beta_1 = k\alpha$, $k \in \mathbb{R}$. 由于 $\beta_2 = \beta - \beta_1$ 与 α 垂直, 我们有

$$0 = (\beta_2, \alpha) = (\beta, \alpha) - k(\alpha, \alpha).$$

这说明 $k = \frac{(\beta, \alpha)}{(\alpha, \alpha)}$, 从而

$$\beta_1 = \frac{(\beta, \alpha)}{(\alpha, \alpha)}\alpha, \quad \beta_2 = \beta - \frac{(\beta, \alpha)}{(\alpha, \alpha)}\alpha.$$

不难直接验证, 这样的分解符合要求, 并且满足要求的分解是唯一的. 这样得到的 k 被称为 β 到向量 α 上的标量投影 (scalar projection), 而 β_1 被称为 β 到向量 α 上的向量投影 (vector projection), 或简称为投影. 而 β_2 将被称作 β 的与 α 正交的分量.

(3) 若前面讨论中的向量 α, β 线性无关, 那么 $\beta_2 = \beta - k\alpha \neq 0$, 从而我们得到了正交向量组 α, β_2 .

(4) 更一般地, 设 $\alpha_1, \dots, \alpha_m$ 是欧氏空间 V 中的一个正交向量组, 对于任意的 $\beta \in V$, 我们可以考虑分解 $\beta = \beta_1 + \beta_2$, 其中

$$\beta_1 = \sum_{i=1}^m \frac{(\beta, \alpha_i)}{(\alpha_i, \alpha_i)}\alpha_i, \quad \beta_2 = \beta - \beta_1.$$

显然, β_1 属于线性子空间 $\langle \alpha_1, \dots, \alpha_m \rangle$, 并且对于任意的 $\gamma := \sum_{i=1}^m k_i \alpha_i \in \langle \alpha_1, \dots, \alpha_m \rangle$, 我们有

$$(\beta_2, \gamma) = \left(\beta - \sum_{i=1}^m \frac{(\beta, \alpha_i)}{(\alpha_i, \alpha_i)}\alpha_i, \sum_{j=1}^m k_j \alpha_j \right)$$

$$\begin{aligned}
&= \sum_{j=1}^m k_j(\beta, \alpha_j) - \underbrace{\sum_{i=1}^m \sum_{j=1}^m k_j \frac{(\beta, \alpha_i)}{(\alpha_i, \alpha_i)} (\alpha_i, \alpha_j)}_{\text{当 } j \neq i \text{ 时 } (\alpha_i, \alpha_j) = 0} \\
&= \sum_{j=1}^m k_j(\beta, \alpha_j) - \sum_{i=1}^m k_i \frac{(\beta, \alpha_i)}{(\alpha_i, \alpha_i)} (\alpha_i, \alpha_i) \\
&= 0.
\end{aligned}$$

这说明上面构造的 β_2 与子空间 $\langle \alpha_1, \dots, \alpha_m \rangle$ 中的所有向量都正交. 由此, 我们称 β_1 为向量 β 到子空间 $\langle \alpha_1, \dots, \alpha_m \rangle$ 的向量投影.

特别地, 若 $\beta \notin \langle \alpha_1, \dots, \alpha_m \rangle$, 那么 $\beta_2 \neq 0$, 从而得到更大的正交向量组 $\alpha_1, \dots, \alpha_m, \beta_2$.

有了上面的准备, 我们首先来看 Gram-Schmidt 正交化过程 (orthogonalization).

定理 7.2.22. 设 $\alpha_1, \alpha_2, \dots, \alpha_s$ 是欧氏空间 V 的一个线性无关的向量组, 令

$$\begin{cases} \beta_1 = \alpha_1, \\ \beta_2 = \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)} \beta_1, \\ \vdots \\ \beta_s = \alpha_s - \sum_{j=1}^{s-1} \frac{(\alpha_s, \beta_j)}{(\beta_j, \beta_j)} \beta_j, \end{cases}$$

则 $\beta_1, \beta_2, \dots, \beta_s$ 是一个正交向量组, 并且与 $\alpha_1, \dots, \alpha_s$ 等价.

证明. 对于 s 作归纳法. $s=1$ 时显然成立. 假设 $s=k$ 时命题成立, 现在来看 $s=k+1$ 的情形. 此时,

$$\beta_{k+1} = \alpha_{k+1} - \sum_{j=1}^k \frac{(\alpha_{k+1}, \beta_j)}{(\beta_j, \beta_j)} \beta_j.$$

由归纳假设可知, 向量组 β_1, \dots, β_k 可由向量组 $\alpha_1, \dots, \alpha_k$ 线性表示. 从而由上面的等式可知, β_{k+1} 可以由 $\alpha_1, \alpha_2, \dots, \alpha_k, \alpha_{k+1}$ 线性表示. 当然, 上面等式的简单变形也指出 α_{k+1} 可以由 $\beta_1, \beta_2, \dots, \beta_k, \beta_{k+1}$ 线性表示. 从而由归纳假设, $\alpha_1, \dots, \alpha_{k+1}$ 与 $\beta_1, \dots, \beta_{k+1}$ 是等价的. 特别地, $\beta_{k+1} \neq 0$.

另外, 在前面 (4) 中的讨论说明, β_{k+1} 与 β_i ($i=1, 2, \dots, k$) 正交. 由归纳假设, 这说明 $\beta_1, \dots, \beta_{k+1}$ 构成一个正交向量组. \square

若将归一化的方法结合在一起, 我们就可以完成定理 7.2.21 的证明了.

定理 7.2.21 的证明. 设 $\alpha_1, \dots, \alpha_n$ 是 n 维欧氏空间 V 的一组基. 由上面的 Gram-Schmidt 正交化过程, 我们得到 V 的一组正交基 β_1, \dots, β_n . 对于每个 i , 令 $\epsilon_i = \frac{\beta_i}{\|\beta_i\|}$. 则 $\epsilon_1, \dots, \epsilon_n$ 构成了 V 的一组标准正交基. \square

注 7.2.23. 若是将 *Gram-Schmidt* 正交化过程与归一化过程结合在一起, 则算法大致如下

$$\begin{array}{ccc}
 \beta_1 = \alpha_1 & \longrightarrow & \epsilon_1 = \beta_1 / \|\beta_1\| \\
 & \swarrow & \\
 \beta_2 = \alpha_2 - (\alpha_2, \epsilon_1)\epsilon_1 & \longrightarrow & \epsilon_2 = \beta_2 / \|\beta_2\| \\
 & \swarrow & \\
 \beta_3 = \alpha_3 - (\alpha_3, \epsilon_1)\epsilon_1 - (\alpha_3, \epsilon_2)\epsilon_2 & \longrightarrow & \epsilon_3 = \beta_3 / \|\beta_3\| \\
 & \swarrow & \\
 \beta_n = \alpha_n - \sum_{j=1}^{n-1} (\alpha_n, \epsilon_j)\epsilon_j & \longrightarrow & \epsilon_n = \beta_n / \|\beta_n\|
 \end{array}$$

例 7.2.24. 用 *Gram-Schmidt* 正交化方法把 \mathbb{R}^3 的一组基

$$\alpha_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$$

化成一组标准正交基.

解.

$$\beta_1 = \alpha_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \xrightarrow{\text{由于 } \|\beta_1\| = \sqrt{3}} \epsilon_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

$$\beta_2 = \alpha_2 - (\alpha_2, \epsilon_1)\epsilon_1 = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} - \frac{2}{\sqrt{3}} \left(\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) = \frac{1}{3} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix},$$

由于 $\|\beta_2\| = \frac{\sqrt{6}}{3}$,

$$\epsilon_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}.$$

$$\begin{aligned}
 \beta_3 &= \alpha_3 - (\alpha_3, \epsilon_1)\epsilon_1 - (\alpha_3, \epsilon_2)\epsilon_2 \\
 &= \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} - \frac{4}{\sqrt{3}} \left(\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) - \frac{2}{\sqrt{6}} \left(\frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \right) = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}.
 \end{aligned}$$

由于 $\|\beta_3\| = 2\sqrt{2}$,

$$\epsilon_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

则 $\epsilon_1, \epsilon_2, \epsilon_3$ 是所需的标准正交基.

□

类似于上面的正交化的计算方法, 必须熟练掌握.