例 7.3.11. **置换矩阵** (*permutation matrix*) 是将单位矩阵的各列重新排列得到的矩阵, 即形如 $(\boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_2}, \ldots, \boldsymbol{e}_{i_n})$ 的矩阵, 其中列向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ 为 $\mathbb{R}^n$ 的自然基, 而 $\sigma = (i_1\ i_2\ \cdots\ i_n) \in \mathcal{S}_n$ 是**一个置换** (也称作**排列**; 请复习教材在定义 *4.3.2* 中给出的关于排列的记号). 显然, 置换矩阵的列向量组仍然构成标准正交基, 从而该矩阵是正交矩阵, 并且, 这样的矩阵是一个第一类正交矩阵的充要条件是其对应的置换 σ 为一个偶置换. 另外, 置换矩阵也可以视为由单位矩阵的各行重新排列得到的矩阵.

例 7.3.12. (1) 设 $\mathscr{A}$ 与 $\mathscr{B}$ 是有限维线性空间 V 上的线性变换, 在基 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ 下的矩阵分别为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$. 不难验证, 映射的复合 $\mathscr{A} \circ \mathscr{B}$ 仍然是 V 上的线性变换, 并且相应的矩阵为 $\boldsymbol{AB}$. 请自行验证这一点

事实上, 若 $V = F^n$ 是数组空间, $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ 是自然基, $\boldsymbol{x} \in V$ 写成列向量形式, 则 $\mathscr{A} : \boldsymbol{x} \mapsto \boldsymbol{Ax}$, $\mathscr{B} : \boldsymbol{x} \mapsto \boldsymbol{Bx}$. 通过直接验证, 我们看到 $\mathscr{A} \circ \mathscr{B} : \boldsymbol{x} \rightsquigarrow \boldsymbol{Bx} \rightsquigarrow \boldsymbol{ABx}$. 这间接表明了 $\mathscr{A} \circ \mathscr{B}$ 对应的矩阵应当为 $\boldsymbol{AB}$.

进一步地, 若 V 是欧氏空间, $\mathscr{A}$ 与 $\mathscr{B}$ 都是正交变换, 则对于任意的 $\boldsymbol{x}, \boldsymbol{y} \in V$,

$$(\mathscr{A}(\mathscr{B}(\boldsymbol{x})), \mathscr{A}(\mathscr{B}(\boldsymbol{y}))) = (\mathscr{B}(\boldsymbol{x}), \mathscr{B}(\boldsymbol{y})) = (\boldsymbol{x}, \boldsymbol{y}).$$

这说明正交变换的复合仍然是正交变换. 这对应于前面提到的事实: 同阶正交矩阵的乘积仍然是正交矩阵.

(2) 设线性变换 $\mathscr{A}$ 是有限维线性空间 V 上的**一一的满射** (也称为**自同构**), 即, V 中的任何元素在 $\mathscr{A}$ 下都存在唯一一个原像. 此时, 不难看出, $\mathscr{A}$ 存在**逆映射** $\mathscr{A}^{-1} : V \to V$, 将 V 中的元素映射成为其在 $\mathscr{A}$ 下的原像. $\mathscr{A}^{-1}$ 仍然是线性映射, 从而是 V 上的线性变换, 称为 $\mathscr{A}$ 的**逆变换**. 进一步地, $\mathscr{A}^{-1}$ 仍然是一一的满射, 满足 $(\mathscr{A}^{-1})^{-1} = \mathscr{A}$.

设 V 是有限维的, 而 $\mathscr{A}^{-1}$ 与 $\mathscr{A}$ 在 V 的基 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ 下的矩阵为 $\boldsymbol{B}$ 和 $\boldsymbol{A}$. 由于 $\mathscr{A}^{-1} \circ \mathscr{A} = \mathrm{id}_V$, 这说明 $\boldsymbol{BA} = \boldsymbol{I}_n$, 从而 $\mathscr{A}^{-1}$ 对应的矩阵 $\boldsymbol{B}$ 恰为 $\boldsymbol{A}^{-1}$.

设 $\mathscr{A}$ 是有限维欧氏空间 V 上的正交变换, 不难验证 $\mathscr{A}$ 存在逆变换 (留作习题). 对于任意的 $\boldsymbol{x}, \boldsymbol{y} \in V$, 由于 $\mathscr{A}^{-1} \circ \mathscr{A} = \mathrm{id}_V = \mathscr{A} \circ \mathscr{A}^{-1}$,

$$(\mathscr{A}^{-1}(\boldsymbol{x}), \mathscr{A}^{-1}(\boldsymbol{y})) \xlongequal{\mathscr{A}\ \text{为正交变换}} (\mathscr{A}(\mathscr{A}^{-1}(\boldsymbol{x})), \mathscr{A}(\mathscr{A}^{-1}(\boldsymbol{y}))) = (\boldsymbol{x}, \boldsymbol{y}),$$

这说明 $\mathscr{A}^{-1}$ 仍然是正交变换. 这对应于前面提到的事实: 正交方阵的逆仍然是正交方阵.

由于我们现在是在实数域上考虑, 线性变换一般不一定有 (实) 特征值.

命题 7.3.13. 设 $\mathscr{A}$ 是 n 维欧氏空间 V 上的正交变换. 教材命题 7.3.2

(1) 若 $\lambda \in \mathbb{R}$ 是 $\mathscr{A}$ 的特征值, 则 $\lambda = \pm 1$.

(2) 若 V 是 n 维欧氏空间, 其中 n 为奇数, 而 $\mathscr{A}$ 为第一类正交变换, 则 $\mathscr{A}$ 一定存在值为 1 的特征值.

(3) 若 V 是 n 维欧氏空间, 而 $\mathscr{A}$ 为第二类正交变换, 则 $\mathscr{A}$ 一定存在值为 -1 的特征值.

证明. (1) $\mathscr{A}$ 在标准正交基下的矩阵为正交方阵, 而我们之前的讨论 (例 6.3.13(4)) 说明了正交矩阵的复特征值 λ 的模长必为 1. 由于线性变换的特征值是它在任何一组基下的矩阵的特征值, 从而 $\lambda = \pm 1$.

注意, 若 V 的维数不一定有限, 但是 $\lambda \in \mathbb{R}$ 是 $\mathscr{A}$ 的特征值, 则我们也可以直接证明 $\lambda = \pm 1$. 证明如下. 设 $\boldsymbol{x} \in V$ 是属于 λ 的特征向量, 从而 $\mathscr{A}(\boldsymbol{x}) = \lambda \boldsymbol{x}$. 此时, 由正交性,

$$(\mathscr{A}(\boldsymbol{x}), \mathscr{A}(\boldsymbol{x})) = (\boldsymbol{x}, \boldsymbol{x}).$$

另一方面, 左式为

$$(\lambda \boldsymbol{x}, \lambda \boldsymbol{x}) = \lambda^2 (\boldsymbol{x}, \boldsymbol{x}).$$

由于 $\boldsymbol{x} \neq \boldsymbol{0}$, 故 $(\boldsymbol{x}, \boldsymbol{x}) \neq 0$. 从而上面两式说明 $\lambda^2 = 1$, 即 $\lambda = \pm 1$.

(2) 特征多项式 $p_{\mathscr{A}}(\lambda)$ 是实系数的 n 次多项式. 由于 n 为奇数, $p_{\mathscr{A}}(\lambda)$ 必有实根. 注意到实系数多项式的虚根必然成对出现, 我们可以假定 $\mathscr{A}$ 的所有特征值为

$$\lambda_1, \overline{\lambda_1}, \lambda_2, \overline{\lambda_2}, \ldots, \lambda_m, \overline{\lambda_m}, \lambda_{m+1}, \ldots, \lambda_{n-m}.$$

其中 $\lambda_i, \overline{\lambda_i}$ 为成对出现的复根, $1 \le i \le m$; 而 $\lambda_{m+1}, \ldots, \lambda_{n-m}$ 为实根, 并且由上面的讨论知其为 ± 1.

注意到每对虚根的乘积 $\lambda_i \overline{\lambda_i} = \|\lambda_i\|^2$ 为正实数, 而所有这些特征值的乘积为 $\mathscr{A}$ 的行列式 1, 这说明这 $n - 2m$ (由于 n 是奇数, 这也是个奇数) 个实根 $\lambda_{m+1}, \ldots, \lambda_{n-m}$ 不可能全为 -1, 从而其中必然存在值为 1 的特征值.

(3) 我们可以仿照上面的思路来证明. 不过, 这儿, 我们换一个思路, 证一个等价的结果: 若 $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ 是一个正交矩阵, 且 $|\boldsymbol{A}| = -1$, 那么 -1 是 $\boldsymbol{A}$ 的特征值. 为此, 我们注意到:

$$|(-1)\boldsymbol{I} - \boldsymbol{A}| = |(-1)\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A} - \boldsymbol{A}| = |-\boldsymbol{A}^{\mathsf{T}} - \boldsymbol{I}| \cdot |\boldsymbol{A}|$$

$$= -|-\boldsymbol{A}-\boldsymbol{I}^{\mathsf{T}}| = -|(-1)\boldsymbol{I}-\boldsymbol{A}|.$$

这说明 $|(-1)\boldsymbol{I}-\boldsymbol{A}|=0$, 从而 -1 为 $\boldsymbol{A}$ 的特征值. □

注 7.3.14. 设 $\mathscr{A}$ 是 n 维欧氏空间 V 上的正交变换, 那么 V 中存在一个标准正交基, 使得 $\mathscr{A}$ 在此基下的矩阵为正交矩阵. 而我们之前的讨论 (例 *6.3.13(4)*) 说明了正交矩阵的特征值 λ 的模长必为 1, 这意味着 $|\operatorname{tr}(\mathscr{A})|\le n$.

注 7.3.15. 如果欧氏空间 V 上的正交变换 $\mathscr{A}$ 有两个不同的 (实) 特征值, 那么 $\mathscr{A}$ 的属于不同特征值的特征向量一定正交. 这是因为这两个特征值只能为 $\lambda_1=1$ 和 $\lambda_2=-1$. 若假定 $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$ 分别是它们的特征向量, 那么

$$(\boldsymbol{x}_1,\boldsymbol{x}_2)=(\mathscr{A}\boldsymbol{x}_1,\mathscr{A}\boldsymbol{x}_2)=(\boldsymbol{x}_1,-\boldsymbol{x}_2)=-(\boldsymbol{x}_1,\boldsymbol{x}_2).$$

因此, 只能有 $(\boldsymbol{x}_1,\boldsymbol{x}_2)=0$, 即 $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$ 正交.

注 7.3.16. 上面命题的一个几何意义: 三维欧氏空间的第一类正交变换, 其必保持一个对称轴不变; 进一步的讨论可以表明, 它是绕该对称轴的旋转变换.

设 $\mathscr{A}$ 是这样的一个第一类正交变换, 则 $\lambda=1$ 是它的一个特征值. 设单位向量 ε_1 是 $\lambda=1$ 的一个特征向量, 则对任意 $k\in\mathbb{R}$, $\mathscr{A}(k\varepsilon_1)=k\mathscr{A}(\varepsilon_1)=k\varepsilon_1$. 这说明 $\mathscr{A}$ 保持直线 $\mathbb{R}\varepsilon_1=\{\,k\varepsilon_1\mid k\in\mathbb{R}\,\}$ 上的向量不变. 利用正交化的方法, 我们可以找到 $\varepsilon_2,\varepsilon_3\in\mathbb{R}^3$ 使得 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 为一组标准正交基.

任取 $\boldsymbol{\alpha}\in\mathbb{R}^3$, 假设它在这组基下的坐标为 $(x,y,z)^{\mathsf{T}}$, 它在 $\varepsilon_2\varepsilon_3$ 平面上的投影 $\boldsymbol{\alpha}'=y\varepsilon_2+z\varepsilon_3$ 与 ε_1 垂直, 从而 $\mathscr{A}(\boldsymbol{\alpha}')$ 与 $\mathscr{A}(\varepsilon_1)=\varepsilon_1$ 垂直. 这说明 $\mathscr{A}$ 将 $\varepsilon_2\varepsilon_3$ 平面上的点仍然映射到该平面上, 从而 $\mathscr{A}$ 局限在该平面上后成为该平面的线性变换, 记作 $\mathscr{A}|_{\varepsilon_2\varepsilon_3}$. 另外需要指出的是, $\mathbb{R}^3$ 的标准内积局限在 $\varepsilon_2\varepsilon_3$ 平面后成为该平面的一个内积, 而 $\varepsilon_2,\varepsilon_3$ 在该内积下是一组标准正交基. 设 $\mathscr{A}|_{\varepsilon_2\varepsilon_3}$ 在基 $\varepsilon_2,\varepsilon_3$ 下的矩阵为 $\boldsymbol{A}'$, 则 $\mathscr{A}$ 在 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵为

$$\boldsymbol{A}=\begin{pmatrix}1 & \\ & \boldsymbol{A}'\end{pmatrix}.$$

由于 $\boldsymbol{A}$ 是行列式为 1 的正交矩阵, 这迫使 $\boldsymbol{A}'$ 是行列式为 1 的正交矩阵, 从而 $\mathscr{A}|_{\varepsilon_2\varepsilon_3}$ 是 $\varepsilon_2\varepsilon_3$ 平面上的第一类正交变换, 即旋转变换. 因此, $\mathscr{A}$ 是 $\mathbb{R}^3$ 的绕 ε_1 轴的旋转.

例 7.3.17. 设 $\mathscr{A}$ 是二维欧氏空间 V 上的正交变换.

(1) 如果 $\mathscr{A}$ 是第一类正交变换, 那么 V 中存在一个标准正交基, 使得 $\mathscr{A}$ 在此基下的矩阵为

$$\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix},\qquad 0\le\theta\le\pi.$$

(2) 如果 $\mathscr{A}$ 是第二类正交变换, 那么 V 中存在一个标准正交基, 使得 $\mathscr{A}$ 在此基下的矩阵为

$$\begin{pmatrix} -1 & \\ & 1 \end{pmatrix}.$$

证明. (1) 由于 $\mathscr{A}$ 为第一类正交变换, 存在 V 中的标准正交基 $\varepsilon_1, \varepsilon_2$, 使得 $\mathscr{A}$ 在这组基下的矩阵为 $\boldsymbol{A}$. 由之前的作业题, 知 $\boldsymbol{A} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$, 其中 $0 \le \theta < 2\pi$. 若 $0 \le \theta \le \pi$, 此即为所求. 若否, 则 $\pi < \theta < 2\pi$. 此时, 不难看出, $\mathscr{A}$ 在标准正交基 $\varepsilon_2, \varepsilon_1$ 下的矩阵为 $\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta') & -\sin(\theta') \\ \sin(\theta') & \cos(\theta') \end{pmatrix}$, 其中 $\theta' = 2\pi - \theta$ 满足 $0 < \theta' < \pi$.

(2) 由于 $\mathscr{A}$ 为第二类正交变换, $\mathscr{A}$ 有一个特征值为 -1. 而 $\mathscr{A}$ 的行列式为 -1, 是其特征值的乘积. 这说明 $\mathscr{A}$ 的另一个特征值为 1. 设 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 为单位向量, 分别是 $\mathscr{A}$ 关于 -1 与 1 的特征向量. 由注 7.3.15 可知, $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$ 正交, 从而构成了 V 的标准正交基. 显然, $\mathscr{A}$ 在这组基下的矩阵为 $\operatorname{diag}(-1, 1)$. □

注 7.3.18. 更一般地, 我们有如下的结果.

(1) 设 $\mathscr{A}$ 是 n 维欧氏空间 V 上的正交变换, 那么 V 中存在一个标准正交基, 使得 $\mathscr{A}$ 在此基下的矩阵为

$$\operatorname{diag}\left\{\lambda_1, \dots, \lambda_r, \begin{pmatrix} \cos(\theta_1) & -\sin(\theta_1) \\ \sin(\theta_1) & \cos(\theta_1) \end{pmatrix}, \dots, \begin{pmatrix} \cos(\theta_m) & -\sin(\theta_m) \\ \sin(\theta_m) & \cos(\theta_m) \end{pmatrix}\right\}, \quad (7.1)$$

其中 $\lambda_i = \pm 1$ $(1 \le i \le r, 0 \le r \le n)$, $0 < \theta_j < \pi$ $(1 \le j \le m, 0 \le m \le \frac{n}{2})$.

(2) n 阶正交矩阵 $\boldsymbol{A}$ 一定正交相似于形如 (7.1) 的分块对角矩阵. (**正交相似**是指存在正交矩阵 $\boldsymbol{T}$ 使得 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 形如给定的矩阵)

对称变换与对称矩阵

定义 7.3.19. 设 $\mathscr{A}$ 是欧氏空间 V 上的线性变换. 若对于任意的 $\boldsymbol{a}, \boldsymbol{b} \in V$ 有 $(\mathscr{A}(\boldsymbol{a}), \boldsymbol{b}) = (\boldsymbol{a}, \mathscr{A}(\boldsymbol{b}))$, 则称 $\mathscr{A}$ 是 V 上的**对称变换** (*symmetric transformation*).

例 7.3.20. (1) 零变换和恒等变换都是对称变换.

(2) 在平面 $\mathbb{R}^2$ 上的旋转变换 $\mathscr{A}_\theta$ 一般不为对称变换.

下面我们讨论对称变换与对称矩阵的关系.

定理 7.3.21. 设 $\mathscr{A}$ 是 n 维欧氏空间 V 上的线性变换, 则以下几条等价:

(1) $\mathscr{A}$ 是 V 上的对称变换;

(2) $\mathscr{A}$ 在任何一组标准正交基下的矩阵都是实对称方阵;

(3) $\mathscr{A}$ 在给定的一组标准正交基下的矩阵是实对称方阵.

证明. (1) $\Rightarrow$ (2): 任取 V 的一组标准正交基 $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n$, 假设 $\mathscr{A}$ 在其上的矩阵为 $\boldsymbol{A} = (a_{ij})_{n\times n}$. 依定义, 这说明对于任意的 i 有

$$\mathscr{A}(\boldsymbol{\varepsilon}_i) = \sum_{k=1}^{n} a_{ki}\boldsymbol{\varepsilon}_k.$$

此时,

$$\begin{aligned}
(\mathscr{A}(\boldsymbol{\varepsilon}_i), \boldsymbol{\varepsilon}_j) &= \left(\sum_{k=1}^{n} a_{ki}\boldsymbol{\varepsilon}_k, \boldsymbol{\varepsilon}_j\right) = \sum_{k=1}^{n} a_{ki}(\boldsymbol{\varepsilon}_k, \boldsymbol{\varepsilon}_j) = \sum_{k=1}^{n} a_{ki}\delta_{kj} = a_{ji}, \\
(\boldsymbol{\varepsilon}_i, \mathscr{A}(\boldsymbol{\varepsilon}_j)) &= \left(\boldsymbol{\varepsilon}_i, \sum_{k=1}^{n} a_{kj}\boldsymbol{\varepsilon}_k\right) = \sum_{k=1}^{n} a_{kj}(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_k) = \sum_{k=1}^{n} a_{kj}\delta_{ik} = a_{ij}.
\end{aligned}$$

由于 $\mathscr{A}$ 为对称变换, 这说明对于任意的 i, j 有 $(\mathscr{A}(\boldsymbol{\varepsilon}_i), \boldsymbol{\varepsilon}_j) = (\boldsymbol{\varepsilon}_i, \mathscr{A}(\boldsymbol{\varepsilon}_j))$, 即 $a_{ij} = a_{ji}$, 从而 $\boldsymbol{A}$ 为对称方阵.

(2) $\Rightarrow$ (3): 这是显然的.

(3) $\Rightarrow$ (1): 设 $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n$ 为 V 上事先给定的一组标准正交基, $\mathscr{A}$ 在其上的矩阵为 $\boldsymbol{A} = (a_{ij})_{n\times n}$. 由假设, $\boldsymbol{A}$ 为对称方阵. 从而由上面的推导可以看到, 对任意的 i, j 有

$$(\mathscr{A}(\boldsymbol{\varepsilon}_i), \boldsymbol{\varepsilon}_j) = (\boldsymbol{\varepsilon}_i, \mathscr{A}(\boldsymbol{\varepsilon}_j)). \tag{7.2}$$

此时, 考虑一般的 $\boldsymbol{a} = \sum_{i=1}^{n} a_i\boldsymbol{\varepsilon}_i, \boldsymbol{b} = \sum_{i=1}^{n} b_i\boldsymbol{\varepsilon}_i \in V$. 直接计算, 有

$$\begin{aligned}
(\mathscr{A}(\boldsymbol{a}), \boldsymbol{b}) &= \left(\mathscr{A}\left(\sum_{i=1} a_i\boldsymbol{\varepsilon}_i\right), \sum_{j=1}^{n} b_j\boldsymbol{\varepsilon}_j\right) = \left(\sum_{i=1}^{n} a_i(\mathscr{A}\boldsymbol{\varepsilon}_i), \sum_{j=1} b_j\boldsymbol{\varepsilon}_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j(\mathscr{A}(\boldsymbol{\varepsilon}_i), \boldsymbol{\varepsilon}_j) \xlongequal{(7.2)} \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j(\boldsymbol{\varepsilon}_i, \mathscr{A}(\boldsymbol{\varepsilon}_j)) \\
&= \cdots = (\boldsymbol{a}, \mathscr{A}(\boldsymbol{b})).
\end{aligned}$$

故 $\mathscr{A}$ 为对称变换. $\square$

注 7.3.22. 接着上面的定理的证明, 设 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 是 V 的另外一组基, $\boldsymbol{\varepsilon}_1,\dots,\boldsymbol{\varepsilon}_n$ 到 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 的过渡矩阵为 $\boldsymbol{P}$, 则 $\mathscr{A}$ 在 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 下的矩阵为 $\boldsymbol{P}^{-1}\boldsymbol{AP}$. 在一般情形下, $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 不再是对称方阵. 另一方面, 若 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 仍然是 V 的一组标准正交基, 则由定理 7.3.21 知 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 必为对称方阵. 事实上, 由命题 *7.3.9* 我们知此时的过渡矩阵 $\boldsymbol{P}$ 为实正交方阵, 故由 $\boldsymbol{A}$ 为对称方阵, 我们可以直接验证 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{P}^{\mathsf{T}}\boldsymbol{AP}$ 为对称方阵.

实对称阵的对角化 一般而言, 实方阵的特征多项式可能有虚根, 从而不能实相似对角化. 但是实对称的矩阵的特征值都是实数.

命题 7.3.23. 设 $\boldsymbol{A}$ 为 n 阶实对称方阵, 则 $\boldsymbol{A}$ 所有的复特征值其实都是实数, 而 $\boldsymbol{A}$ 的属于不同特征值的实特征向量在 $\mathbb{R}^n$ 的标准内积下必然正交.

证明. 设复数 λ 是 $\boldsymbol{A}$ 的特征值, 从而存在非零复向量 $\boldsymbol{x}\in\mathbb{C}^n$, 使得 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$. 对该式两边取复共轭有 $\overline{\boldsymbol{Ax}}=\overline{\lambda\boldsymbol{x}}$, 即 $\overline{\boldsymbol{A}}\overline{\boldsymbol{x}}=\overline{\lambda}\overline{\boldsymbol{x}}$. 由于 $\boldsymbol{A}$ 为实对称矩阵, 即 $\overline{\boldsymbol{A}}=\boldsymbol{A}=\boldsymbol{A}^{\mathsf{T}}$. 这说明

$$\lambda\overline{\boldsymbol{x}}^{\mathsf{T}}\boldsymbol{x}=\overline{\boldsymbol{x}}^{\mathsf{T}}(\lambda\boldsymbol{x})=\overline{\boldsymbol{x}}^{\mathsf{T}}(\boldsymbol{AX})=(\overline{\boldsymbol{x}}^{\mathsf{T}}\overline{\boldsymbol{A}}^{\mathsf{T}})\boldsymbol{X}=(\overline{\boldsymbol{Ax}})^{\mathsf{T}}\boldsymbol{x}=(\overline{\lambda\boldsymbol{x}})^{\mathsf{T}}\boldsymbol{x}=\overline{\lambda}\overline{\boldsymbol{x}}^{\mathsf{T}}\boldsymbol{x}.$$

这说明 $(\lambda-\overline{\lambda})\overline{\boldsymbol{x}}^{\mathsf{T}}\boldsymbol{x}=0$. 由于 $\boldsymbol{x}\neq\boldsymbol{0}$, 我们有 $\overline{\boldsymbol{x}}^{\mathsf{T}}\boldsymbol{x}=\|\boldsymbol{x}\|^2\neq0$, 从而 $\lambda=\overline{\lambda}$, 即 $\lambda\in\mathbb{R}$.

接下来设 $\lambda_1,\lambda_2\in\mathbb{R}$ 为 $\boldsymbol{A}$ 不同的特征值, 而 $\boldsymbol{x}_1,\boldsymbol{x}_2\in\mathbb{R}^n$ 为相应的特征向量. 于是有

$$\lambda_1\boldsymbol{x}_1^{\mathsf{T}}\boldsymbol{x}_2=(\lambda_1\boldsymbol{x}_1)^{\mathsf{T}}\boldsymbol{x}_2=(\boldsymbol{Ax}_1)^{\mathsf{T}}\boldsymbol{x}_2=\boldsymbol{x}_1^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}}\boldsymbol{x}_2=\boldsymbol{x}_1^{\mathsf{T}}(\boldsymbol{Ax}_2)=\boldsymbol{x}_1^{\mathsf{T}}(\lambda_2\boldsymbol{x}_2)=\lambda_2\boldsymbol{x}_1^{\mathsf{T}}\boldsymbol{x}_2.$$

由于 $\lambda_1\neq\lambda_2$, 故有 $\boldsymbol{x}_1^{\mathsf{T}}\boldsymbol{x}_2=0$. □

对称变换也有相应的性质.

定理 7.3.24. 设 $\mathscr{A}$ 是有限维欧氏空间 V 上的对称变换, 则 $\mathscr{A}$ 的特征值都是实数, 且 $\mathscr{A}$ 在不同特征值下的特征向量相互正交.

证明. 关于第一部分断言, 我们先取 V 的一组标准正交基, 并设 $\mathscr{A}$ 在这组基下的矩阵为 $\boldsymbol{A}$. 于是, $\boldsymbol{A}$ 为实对称阵, 而 $\boldsymbol{A}$ 的特征值就是 $\mathscr{A}$ 的特征值. 于是利用命题 7.3.23 即可.

关于第二部分断言, 我们设 V 中的向量 $\boldsymbol{x}_1$ 是属于 λ_1 的特征值, $\boldsymbol{x}_2$ 是属于 λ_2 的特征值, 并且 $\lambda_1\neq\lambda_2$. 此时,

注意这一部分证明在 V 不是有限维时也成立

$$\lambda_1(\boldsymbol{x}_1,\boldsymbol{x}_2)=(\lambda_1\boldsymbol{x}_1,\boldsymbol{x}_2)=(\mathscr{A}\boldsymbol{x}_1,\boldsymbol{x}_2)=(\boldsymbol{x}_1,\mathscr{A}\boldsymbol{x}_2)=(\boldsymbol{x}_1,\lambda_2\boldsymbol{x}_2)=\lambda_2(\boldsymbol{x}_1,\boldsymbol{x}_2).$$

由于 $\lambda_1\neq\lambda_2$, 该式说明 $(\boldsymbol{x}_1,\boldsymbol{x}_2)=0$, 从而 $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$ 正交. □

引理 7.3.25. *如果 n 阶实矩阵 $\boldsymbol{A}$ 的特征多项式的复根都是实数, 那么 $\boldsymbol{A}$ 一定可以正交相似上三角化, 即存在正交矩阵 $\boldsymbol{Q}$ 使得 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ 为上三角阵.*

证明. 由于实对称阵 $\boldsymbol{A}$ 的特征多项式的复根都是实数, 利用 Schur 定理 (定理 6.4.22) 及其随后的注, 我们知道 $\boldsymbol{A}$ 可以通过实矩阵相似上三角化, 即存在可逆的实矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为实的上三角阵. 由定理 7.2.25 可知, 实的可逆矩阵 $\boldsymbol{P}$ 存在 QR 分解: $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{R}$, 其中 $\boldsymbol{Q}$ 为正交矩阵, $\boldsymbol{R}$ 为上三角矩阵. 此时, $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{R}\boldsymbol{B}\boldsymbol{R}^{-1}$ 为上三角阵. $\square$