线性代数: 特征向量与特征值

Linear algebra: Eigenvectors and eigenvalues

王浩铭

2017 年 · 冬

这篇笔记的参考资料为刘丽、韩本三《高等代数》，部分内容根据我的理解进行调整。因为本人水平有限，无法保证本文内容正确性，这篇笔记仅供参考。若您发现本文的错误，请将这些错误发送到我的邮箱 wanghaoming17@163.com ，谢谢！您可以在我的主页中浏览更多笔记。

目录

1 特征值与特征向量的概念与计算

1.1 特征值与特征向量概念

定义 1.1. 设 $A = A_{n\times n}$ 是数域 P 上的 n 阶方阵，若在数域 P 中存在数 λ 和非零 n 维列向量 X，使得

$$AX = \lambda X,$$

则称 λ 为矩阵 A 的特征值，称 X 为 A 属于 λ 的特征向量.

Edition 4

如何计算矩阵的特征值呢，易知 $AX=\lambda X$ 等价于 $(\lambda E-A)X=\theta$，只需求出该方程组的解 X 代入便可求得 λ，这便与上一章的内容联系了起来.

易知矩阵 A 有特征值与特征向量的充要条件为方程组

$$(\lambda E-A)X=\theta$$

有非零解，而齐次线性方程组有非零解的充要条件为行列式 $|\lambda E-A|=0$，因此 λ_0 是矩阵 A 的特征值的充要条件为满足 $|\lambda_0 E-A|=0$

根据行列式的定义，在行列式

$$|\lambda E-A|=\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}$$

的展开式中，有一项是主对角元的连乘积

$$(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn})$$

而其余各项中最多含义 $n-2$ 个主对角元，从而 λ 的最高次为 $n-2$，因此 $|\lambda_0 E-A|$ 的展开式中 λ 的 n 次与 $n-1$ 次项只能在主对角元的连乘积中出现，分别为

$$\lambda^n,\ -(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1},$$

令 $\lambda=0$，则 $|\lambda_0 E-A|=|-A|=(-1)^n|A|$，因此 $|\lambda_0 E-A|$ 是 λ 的 n 次多项式，令其为 $f(\lambda)$，则

$$f(\lambda)=|\lambda_0 E-A|=\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^n|A|, \tag{1}$$

由此，矩阵特征值的问题转化为求 n 次多项式 $f(\lambda)$ 的根的问题.

定义 1.2. *设 $A=A_{n\times n}$ 是数域 P 上的 n 阶方阵，矩阵 $\lambda E-A$ 称为 A 的特征矩阵，$|\lambda E-A|$ 称为 A 的特征多项式，$|\lambda E-A|=0$ 称为 A 的特征方程，特征方程的解称为 A 的特征根，也就是特征值.*

根据代数基本定理，在复数域内，n 次方程恰好有 n 个根（k 重根算 k 个根），因此在复数域内，n 次方程 $|\lambda E-A|=0$ 恰好有 n 个特征值，且属于特征值 λ_0 的特征向量正是方程组

$$(\lambda_0 E-A)X=\theta$$

的全体非零解向量，由此给出计算方阵 A 的特征值与特征向量的步骤：

1. 计算 n 阶矩阵 A 的特征方程 $|\lambda E-A|=0$ 的 n 个解，这是矩阵 A 的全部特征值；

2. 设 $\lambda_1,\lambda_2,...,\lambda_s$ 是 A 的全部互异特征值，分别把每个 λ_i 代入方程组 $(\lambda_i E-A)X=\theta$ 中，求得基础解系，该基础解系的线性组合（组合系数不全为 0）正是 A 属于特征值 λ_i 的全体特征向量.

1.2 特征值与特征向量的性质

定理 1.1 (特征值与特征向量的基本运算性质)**.** 特征值与特征向量有如下基本运算性质:

1. 若 X 是 A 属于 λ 的特征向量, 则 $kX(k \neq 0)$ 也是 A 属于 λ 的特征向量;
2. 若 $X_1, X_2, ..., X_s$ 是 A 属于 λ 的特征向量, 则它们的线性组合 $k_1X_1 + k_2X_2 + \cdots + k_sX_s$ 也是 A 属于 λ 的特征向量;
3. 设 $|A| \neq 0, X \neq \theta$ 且 $AX = \lambda X$, 则 $\frac{1}{\lambda}$ 是 A^{-1} 的一个特征值, 且 X 是 A^{-1} 属于特征值 $\frac{1}{\lambda}$ 的特征向量;
4. 设 $|A| \neq 0, X \neq \theta$ 且 $AX = \lambda X$, 则 $\frac{|A|}{\lambda}$ 是 A^* 的一个特征值, 且 X 是 A^* 属于特征值 $\frac{|A|}{\lambda}$ 的特征向量;
5. 设 λ 是 n 阶矩阵 A 的一个特征值, X 是 A 属于特征值 λ 的特征向量, $f(x) = a_mx^m + a_{m-1}x^{m-1} + \cdots + a_1x + a_0$ 是一个多项式, 则 $f(\lambda)$ 是 $f(A)$ 的一个特征值, 且 X 是 $f(A)$ 属于特征值 $f(\lambda)$ 的特征向量.
6. n 阶矩阵 A 与它的转置矩阵 A^T 有相同的特征值.

证明. 1.2. 的证明略.

3. 因为 A 可逆, 则 $AX = \lambda X$ 两侧左乘 A^{-1} 有
$$A^{-1}AX = \lambda A^{-1}X \Rightarrow \frac{1}{\lambda}X = A^{-1}X;$$
4. 因为 $A^* = |A|A^{-1}$ 因此对 $\frac{1}{\lambda}X = A^{-1}X$ 两侧乘 $|A|$ 有
$$\frac{|A|}{\lambda}X = |A|A^{-1}X = A^*X,$$
5. 利用数学归纳法. 因为 $A^2X = \lambda AX = \lambda^2X$, 设 $A^kX = \lambda^kX$, 因此
$$A^{k+1}X = AA^kX = \lambda^kAX = \lambda^{k+1}X,$$
因此有 $A^mX = \lambda^mX$, 即
$$\begin{aligned} f(A)X &= (a_mA^m + a_{m-1}A^{m-1} + \cdots + a_1A + a_0)X \\ &= a_mA^mX + a_{m-1}A^{m-1}X + \cdots + a_1AX + a_0X \\ &= a_m\lambda^mX + a_{m-1}\lambda^{m-1}X + \cdots + a_1\lambda X + a_0X \\ &= (a_m\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0)X \\ &= f(\lambda)X. \end{aligned}$$
6. 因为
$$|\lambda E - A| = |(\lambda E - A)^T| = |\lambda E - A^T|$$
即 A 与 A^T 有相同的特征多项式, 因此有相同的特征值.

注 1.1. A 与 A^T 有相同的特征值，但对于同一特征值的特征向量不一定相同.

注 1.2. 相同的**特征多项式**是证明两矩阵特征值相同的**重要技巧**，在相似矩阵具有相同特征值中也应用的这一方法.

注 1.3 (应用强调)**.** 对于上述性质，所如下应试型强调：设 A 的特征值为 λ_A

- 若 $B = k_1A + k_2E$，则 B 的特征值为 $\lambda_B = k_1 \cdot \lambda_A + k_2$；
- 若 $C = k_3(A^*)^2 + k_4E$，则 C 的特征值为 $\lambda_C = k_3 \cdot \frac{|A|}{\lambda_A} + k_4$，若知道 A 的全部特征值 λ_{A_i}，则 $|A| = \prod \lambda_{A_i}$；
- 由于若 X 是 A 属于 λ_A 的特征值，则 X 也是 A^{-1}, A^* 属于 $\frac{1}{\lambda_A}, \frac{|A|}{\lambda_A}$ 的特征值，因此有
$$(A^* + A^{-1})X = A^*X + A^{-1}X = \frac{|A|}{\lambda_A}X + \frac{1}{\lambda_A}X = (\frac{|A|}{\lambda_A} + \frac{1}{\lambda_A})X,$$
即若 f_1, f_2, f_3 为多项式，则 $f_1(A) + f_2(A^*) + f_3(A^{-1})$ 的特征值为
$$f_1(\lambda_A) + f_2\left(\frac{|A|}{\lambda_A}\right) + f_3\left(\frac{|A|}{\lambda_A}\right).$$

□

定理 1.2. n 阶矩阵 A 的对角线元素为 a_{ii}，则有

1. $\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}$；
2. $\lambda_1\lambda_2\cdots\lambda_n = |A|$.

证明. 设 n 阶矩阵 A 的特征值分别为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，由代数基本定理可知 A 的特征多项式为

$$\begin{aligned} f(\lambda) &= (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) \\ &= \lambda^n + -(\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots + (-1)\lambda_1\lambda_2\cdots\lambda_n, \end{aligned}$$

比较公式1 可知 $\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}$ 以及 $\lambda_1\lambda_2\cdots\lambda_n = |A|$. □

注 1.4. $x = x_1, x_2, ..., x_m$ 是多项式 $f(x)$ 的根 $\Leftrightarrow f(x) = (x - x_1)(x - x_2)\cdots(x - x_m)$.

推论 1.1. 设 A 是 n 阶矩阵，则 A 可逆的充要条件为 A 没有 0 特征值.

定义 1.3 (迹)**.** n 阶矩阵 A 的主对角线上 n 个元素的和 $a_{11} + a_{22} + \cdots + a_{nn}$ 称为 A 的迹，记作 $\mathrm{tr}(A)$.

定理 1.3. 若 $\lambda_1, \lambda_2, ..., \lambda_m$ 是矩阵 A 的 m 个互异特征值，且 $X_1, X_2, ..., X_m$ 是 A 属于 $\lambda_1, \lambda_2, ..., \lambda_m$ 的特征向量，则 $X_1, X_2, ..., X_m$ 线性无关.

证明. 利用数学归纳法，首先证明 A 的两个互异特征值对应的特征向量线性无关，设 $AX_1 = \lambda_1 X_1$ 以及 $AX_2 = \lambda_2 X_2$，其中 $\lambda_1 \neq \lambda_2$.

利用反证法，设 X_1, X_2 线性相关在，即存在常数 k 使得 $X_1 = kX_2$，因此

$$AX_1 = \lambda_1 X_1 \Rightarrow AkX_1 = \lambda_1 kX_2 \Rightarrow AX_2 = \lambda_1 X_2.$$

因此 $\lambda_1 = \lambda_2$，矛盾，即 X_1, X_2 线性无关.

设对于 A 的 $s-1$ 个互异特征值对应的特征向量线性无关，则对于 s 个互异特征值 $\lambda_1, \lambda_2, ..., \lambda_s$ 对应的特征向量线性 $X_1, X_2, ..., X_s$ 可以利用反证法证明其线性无关：

假设 $X_1, X_2, ..., X_s$ 线性相关，由定理??可知向量组 $(X_1, X_2, ..., X_s)$ 中存在向量 X_i 使得 X_i 可以被其余 $s-1$ 个向量线性表示，因为 $X_i \neq \theta$，因此即存在一组不全为零的常数 $c_1, ..., c_{i-1}, c_{i+1}, ..., c_s$ 使得

$$X_i = c_1 X_1 + ... + c_{i-1} X_{i-1} + c_{i+1} X_{i+1} + ... + c_s X_s,$$

代入

$$AX_i = \lambda_i X_i$$

有

$$\begin{aligned}
&A(c_1 X_1 + ... + c_{i-1} X_{i-1} + c_{i+1} X_{i+1} + ... + c_s X_s) \\
&= c_1 AX_1 + ... + c_{i-1} AX_{i-1} + c_{i+1} AX_{i+1} + ... + c_s AX_s \\
&= c_1 \lambda_1 X_1 + ... + c_{i-1} \lambda_{i-1} X_{i-1} + c_{i+1} \lambda_{i+1} X_{i+1} + ... + c_1 \lambda_s X_s \\
&= \lambda_i (c_1 X_1 + ... + c_{i-1} X_{i-1} + c_{i+1} X_{i+1} + ... + c_s X_s) \\
&= c_1 \lambda_i X_1 + ... + c_{i-1} \lambda_i X_{i-1} + c_{i+1} \lambda_i X_{i+1} + ... + c_1 \lambda_i X_s,
\end{aligned}$$

因此有

$$\begin{aligned}
&c_1 \lambda_1 X_1 + ... + c_{i-1} \lambda_{i-1} X_{i-1} + c_{i+1} \lambda_{i+1} X_{i+1} + ... + c_1 \lambda_s X_s \\
&= c_1 \lambda_i X_1 + ... + c_{i-1} \lambda_i X_{i-1} + c_{i+1} \lambda_i X_{i+1} + ... + c_1 \lambda_i X_s,
\end{aligned}$$

以及

$$c_1(\lambda_1 - \lambda_i) X_1 + ..., c_{i-1}(\lambda_{i-1} - \lambda_i) X_{i-1} + c_{i+1}(\lambda_{i+1} - \lambda_i) X_{i+1} + ... + c_s(\lambda_s - \lambda_i) X_s = \theta.$$

由假设，A 的 $s-1$ 个互异特征值对应的特征向量线性无关，且 $c_1, ..., c_{i-1}, c_{i+1}, ..., c_s$ 不全为零，设 $c_t \neq 0$，则必有 $\lambda_i = \lambda_t$，矛盾，因此 A 的任意 s 个互异特征值对应的特征向量线性无关. □

推论 1.2. *若 $\lambda_1, \lambda_2, ..., \lambda_m$ 是矩阵 A 的 m 个互异特征值，且 $X_{i1}, X_{i2}, ..., X_{im}$ 是 A 属于 λ_i 的线性无关特征向量，则 $X_{11}, X_{12}, ..., X_{1r_1}, X_{21}, X_{22}, ..., X_{2r_2}, ..., X_{m1}, X_{m2}, ..., X_{mr_m}$ 线性无关.*

证明. 由定理1.3可知，对于任意 j_i 有 $X_{1j_1}, X_{2j_2}, ..., X_{mj_m}$ 线性无关，又因为 $X_{i1}, X_{i2}, ..., X_{im}$ 线性无关，因此 $X_{11}, X_{12}, ..., X_{1r_1}, X_{21}, X_{22}, ..., X_{2r_2}, ..., X_{m1}, X_{m2}, ..., X_{mr_m}$ 线性无关. □

定理 1.4. *若 λ_0 是 n 阶矩阵的 k 重特征值，则 A 属于 λ_0 的线性无关的特征向量最多有 k 个.*

2 相似矩阵与矩阵的相似对角化

2.1 相似矩阵及其性质

定义 2.1 (相似)**.** 设 A, B 为 n 阶矩阵，若存在可逆矩阵 P，使得

$$B = P^{-1}AP$$

则称 A 与 B 相似，记为 $A \backsim B$，并称 P 为相似变换矩阵.

定理 2.1. 矩阵的详细关系满足三条基本性质：

1. 反身性：$A \backsim A$;
2. 对称性：$A \backsim B \Rightarrow B \backsim A$;
3. 传递性：$A \backsim B, B \backsim C \Rightarrow A \backsim C$.

证明. 1. 因为 $A = EAE = E^{-1}AE$，所以 $A \backsim A$;

2. 令 $P^{-1} = Q$, 则易知 Q 可逆, 且 $Q^{-1} = P$, 因为 $B = P^{-1}AP$ 所以 $PBP^{-1} = PP^{-1}APP^{-1} = A$，即 $Q^{-1}BQ = A$，因此 $B \backsim A$;

3. 设 $P^{-1}AP = B, Q^{-1}BQ = C$，因此 $Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ) = C$，因此 $A \backsim C$.

□

定理 2.2. 设 A, B 为 n 阶矩阵，且 $A \backsim B$，则

1. $R(A) = R(B)$;
2. A 与 B 有相同的特征值;
3. $|A| = |B|$
4. $\mathrm{tr}(A) = \mathrm{tr}(B)$;
5. A 与 B 有相同的可逆性，若均可逆则 $A^{-1} \backsim B^{-1}$;
6. 设 $f(x)$ 为多项式，则 $f(A) \backsim f(B)$;
7. 若 $A_i \backsim B_i, i = 1, 2, ..., s$，则 $\mathrm{diag}(A_1, A_2, ..., A_s) \backsim \mathrm{diag}(B_1, B_2, ..., B_s)$.

证明. 1. 因为 $B = P^{-1}AP$，因此 $R(B) = R(P^{-1}AP) \leq R(A)$，同理 $A = PBP^{-1}$，因此 $R(A) = R(PBP^{-1}) \leq R(B)$，因此 $R(A) = R(B)$;

注 2.1. $R(P^{-1}AP) \leq \min\{P^{-1}, AP\} \leq R(AP) \leq \min\{R(A), R(P)\} \leq R(A)$.

2. 因为

$$|\lambda E - B| = |\lambda E - P^{-1}AP| = |\lambda P^{-1}P - P^{-1}AP| = |P^{-1}| \cdot |\lambda E - A| \cdot |P| = |\lambda E - A|.$$

因此 A 与 B 有相同的特征多项式，即 A 与 B 有相同的特征值;

3. 因为 A 与 B 有相同的特征值 $\lambda_1, \lambda_2, ..., \lambda_n$，因此 $|A| = |B| = \lambda_1\lambda_2\cdots\lambda_n$；

4. 因为 A 与 B 有相同的特征值 $\lambda_1, \lambda_2, ..., \lambda_n$，因此 $\mathrm{tr}(A) = \mathrm{tr}(B) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$；

5. 因为 $|A| = |B|$，因此 A, B 有相同的可逆性，因为
$$E = BB^{-1} = P^{-1}APB^{-1} \Rightarrow P^{-1}A^{-1}P = B^{-1};$$

6. 因为 $B = P^{-1}AP$，因此
$$B^m = P^{-1}APP^{-1}AP\cdots P^{-1}AP = P^{-1}A^mP,$$
因此有 $P^{-1}f(A)P = f(B)$;

7. 设 $P_i^{-1}A_iP_i = B_i$，则构造矩阵 $P = \mathrm{diag}(P_1, P_2, ..., P_s)$，则 $P^{-1} = \mathrm{diag}(P_1^{-1}, P_2^{-1}, ..., P_s^{-1})$，因此 $P^{-1}AP = B$.

□

注 2.2. 矩阵 A, B 有相同特征值是两者相似的必要条件，但若 A 的特征值均为单根，且与 B 的特征值相同，则是两者相似的充分条件；或者 A, B 非互异特征值线性无关特征向量的个数等于重根充数，也可以推出两者相似，因为它们都相似于对角阵.

2.2 n 阶方阵 A 的对角化

若矩阵与对角阵相似，则称其可以**对角化**.

定理 2.3 (可对角化的充要条件). n 阶矩阵 A 与对角矩阵 Λ 相似的充要条件为 A 有 n 个线性无关的特征向量.

证明. 必要性：设 $A \backsim \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$，则存在可逆矩阵 P 使得
$$P^{-1}AP = \Lambda,$$
对 P 按列分块有 $P = (X_1, X_2, ..., X_n)$，因为 P 可逆，因此 $X_1, X_2, ..., X_n$ 线性无关，则
$$AP = A(X_1, X_2, ..., X_n) = P\Lambda = (X_1, X_2, ..., X_n)\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$
即
$$(AX_1, AX_2, ..., AX_n) = (\lambda_1X_1, \lambda_2X_2, ..., \lambda_nX_n),$$
所以 $X_1, X_2, ..., X_n$ 是 A 分别属于特征值 $\lambda_1, \lambda_2, ..., \lambda_n$ 的特征向量.

充分性，设 $X_1, X_2, ..., X_n$ 是 A 分别属于特征值 $\lambda_1, \lambda_2, ..., \lambda_n$ 的线性无关的特征向量，即
$$(AX_1, AX_2, ..., AX_n) = (\lambda_1X_1, \lambda_2X_2, ..., \lambda_nX_n),$$

令 $P=(X_1,X_2,...,X_n),\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,...,\lambda_n)$，则

$$AP=P\Lambda,$$

因为 $X_1,X_2,...,X_n$ 线性无关，所以 P 可逆，即

$$P^{-1}AP=\Lambda.$$

□

推论 2.1. *1. 对角阵 Λ 的主对角元恰是 A 的 n 个特征值;*

2. P 的 n 个列向量正是 A 分别属于特征值的线性无关的特征向量;

3. 特征向量在 P 中的列数与特征值在 Λ 中的列数相对应.

定理 2.4. *若 n 阶方阵 A 的特征值都是单根，则 A 与对角阵相似.*

证明. 由定理1.3以及定理2.3可证. □

定理 2.5. *n 阶方阵 A 与对角阵相似的充要条件为 A 的每一个 k 重特征值 λ 恰好对应 k 个线性无关的特征向量.*

注 2.3. 这一定理告诉我们：$A\backsim\Lambda\Leftrightarrow R(\lambda_iE-A)=n-n_i$，其中 λ_i 是 A 的 n_i 重特征根，即 $R(\ker(\lambda_iE-A))=n_i$.

这说明，如果 A 可相似对角化（如 A 是实对称矩阵），且 $R(\lambda_iE-A)=n-1$ 则 λ_i 是 A 单根特征值；一般的，若 $R(\lambda_iE-A)=n-k$，则 λ_j 是 A 的 k 重特征值.

2.3 若尔当标准型

定义 2.2 (若尔当（Jordan）块). 形如

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

的方阵称为若尔当块.

定义 2.3 (若尔当型矩阵). 由若尔当块构成的矩阵

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_n \end{bmatrix}$$

称为若尔当型矩阵，其中 $J_i,i=1,2,...,n$ 为若尔当块.

例 2.1. 例如

$$\begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 1 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 1 \\ 0 & 0 & 0 & 0 & \sqrt{2} \end{bmatrix}, \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & -2 \end{bmatrix}$$

都是若尔当型矩阵.

注 2.4. 一阶矩阵, 即一个数也是若尔当块; 对角阵也是若尔当型矩阵.

定理 2.6. 任意 n 阶复矩阵 A 必相似于一个若尔当型矩阵, 该若尔当型矩阵的主对角元是 A 的全部特征值, 且主对角元是 λ_i 的若尔当块的个数为 $n - R(\lambda_i - A)$.

定义 2.4. 与矩阵 A 相似的若尔当型矩阵称为 A 的若尔当标准形.

3 实对称矩阵的相似对角化

3.1 实对称矩阵的特征值与特征向量的性质

引理 3.1 (共轭复数运算性质)**.** 设 x_1, x_2 为复数, 则 $\overline{x_1 x_2} = \bar{x_1}\bar{x_2}$.

证明. 设 $x_1 = ax + bi, x_2 = cy + di$, 则有 $x_1x_2 = (ax+bi)(cy+di) = acxy + adxi + bcyi + bdi^2 = (acxy - bd) + (adx + bcy)i$; 因此 $\overline{x_1x_2} = (acxy - bd) - (adx + bcy)i$, 则令 $b = -b, d = -d$ 有 $\bar{x_1}\bar{x_2} = (acxy - bd) + (-adx - bcy)i = (acxy - bd) - (adx + bcy)i = \overline{x_1x_2}$. □

定理 3.1. 实对称矩阵的特征值都是实数.

证明. 设 A 为 n 维实对称矩阵, λ 为其特征值, $X = (x_1, x_2, ..., x_n)^T$ 是对于的特征向量, 其中 $x_j = a_j + b_j i, j = 1, 2, ..., n$.

则有

$$AX = \lambda X,$$

两侧取共轭有

$$\overline{AX} = \overline{AX}$$

由引理3.1, 因为 A 为实对称矩阵有

$$\bar{A}\bar{X} = \bar{\lambda}\bar{X} = A\bar{X},$$

两侧左乘 X^T, 因为 A 为对称阵有

$$\bar{\lambda}X^T\bar{X} = X^TA\bar{X} = (AX)^T\bar{X} = \lambda X^T\bar{X},$$

即

$$(\bar{\lambda} - \lambda)X^T\bar{X} = \theta.$$

因为 $X \neq \theta$, 且 $X^T\bar{X} = \sum_{j=1}^n (a_j + b_j i)(a_j - b_j i) = \sum_{j=1}^n (a_j^2 + b_j^2) \neq 0$, 因此 $\lambda = \bar{\lambda}$. □

注 3.1. 一般实 n 阶方阵的特征值不一定为实数.

定理 3.2. 实对称矩阵的不同特征值对应的特征向量正交.

证明. 设 $\lambda_1 \neq \lambda_2$ 是 A 的特征值，X_1, X_2 是对于的特征向量，则

$$AX_1 = \lambda_1 X_1 \Rightarrow X_2^T AX_1 = (AX_2)^T = \lambda_2 X_2^T X_1 = \lambda_1 X_2^T X_1,$$

因此

$$(\lambda_2 - \lambda_1)X_2^T X_1 = 0,$$

因为 $\lambda_1 \neq \lambda_2$，所以 $X_2^T X_1 = 0$ 即 $X_1 \perp X_2$. □

3.2 实对称矩阵对角化的方法

定理 3.3. 任意一个 n 阶实对称矩阵 A，都存在一个 n 阶正交矩阵 Q，使得 $Q^T AQ = Q^{-1}AQ$ 为对角阵.

推论 3.1. 实对称矩阵 A 属于 k 重特征值 λ_0 的线性无关的特征向量正好有 k 个.

注 3.2. 串联定理*1.4*，定理*2.5*，定理*3.1*.

定理3.3说明：

1. 实对称矩阵必与对角阵相似；
2. 实对称矩阵对角化相似变换矩阵可以是 正交矩阵.

回忆**格拉姆-施密特正交化法**，由线性无关的向量组 $(\alpha_1, \alpha_2, ..., \alpha_s)$ 得到的正交向量组 $(\beta_1, \beta_2, ..., \beta_s)$ 中的每一个向量 β_i 都是 $\alpha_1, \alpha_2, ..., \alpha_s$ 的线性组合.

设 $\alpha_1, \alpha_2, ..., \alpha_k$ 为实对称矩阵 A 属于特征值 λ_0 的线性无关的特征向量，即 $\alpha_1, \alpha_2, ..., \alpha_r$ 是 $(\lambda_0 E - A)X = 0$ 的解，则由齐次线性方程组解的性质可知，它们的线性组合仍为 $(\lambda_0 E - A)X = 0$ 的解，即由施密特正交化法得到的正交向量 $\beta_1, \beta_2, ..., \beta_k$ 也是实对称矩阵 A 属于特征值 λ_0 的线性无关的特征向量.

由于实对称矩阵的不同特征值对应的特征向量正交(定理3.2)可求出 n 维实对称矩阵 A 的 n 个两两正交的特征向量，通过单位化即可得到对应的单位正交向量组，由此形成的矩阵正是正交矩阵 Q(定理**??**)，由于 Q 的列向量是其特征向量，因此由可对角化的充要条件(2.3)可知 $Q^T AQ = Q^{-1}AQ = \Lambda$.

总结上述内容，得到求正交矩阵 Q 的步骤：

1. 求出 n 阶实对称矩阵 A 的全部特征值，设 $\lambda_1, \lambda_2, ..., \lambda_s$ 是 A 的全部互异特征值；
2. 对每一个 λ_i，求解对应的线性无关的特征向量组，即求解 $(\lambda_i E - A)X = \theta$ 的基础解析；
3. 若 λ_i 为 n_i 重根，则将求解的基础解析先正交化，在单位化，若 λ_i 为单根，则只进行单位化，得到 A 的 n 个两两正交的特征向量 $\epsilon_{11}, \epsilon_{12}, ..., \epsilon_{1n_1}, \epsilon_{21}, \epsilon_{22}, ..., \epsilon_{2n_2}, ..., \epsilon_{s1}, \epsilon_{s2}, ..., \epsilon_{sn_s}$；
4. 令 $Q = (\epsilon_{11}, \epsilon_{12}, ..., \epsilon_{1n_1}, \epsilon_{21}, \epsilon_{22}, ..., \epsilon_{2n_2}, ..., \epsilon_{s1}, \epsilon_{s2}, ..., \epsilon_{sn_s})$，则 Q 为正交向量组，且满足

$$Q^{-1}AQ = \mathrm{diag}(\underbrace{\lambda_1, ..., \lambda_1}_{n_1}, \underbrace{\lambda_2, ..., \lambda_2}_{n_2}, ..., \underbrace{\lambda_s, ..., \lambda_s}_{n_s}).$$