

广义特征值与酉相似

李立斌 杨巧林

(扬州大学工学院, 扬州, 225001)

摘 要 关于矩阵的特征值及相似理论研究相当深刻, 但对广义特征值及相似问题有关文献很少论述。本文讨论这方面的问题, 它在振动理论、物理及其它领域有着广泛的应用。本文所述数域范围为复数域, 如无特指, 矩阵均为 n 阶复矩阵, 向量为 n 维复向量。

关键词 广义特征值 广义酉相似 瑞利(Rayleigh)商

分类号 O241.6

1 广义特征值

以下恒设 M 为正定埃米特(Hermite)矩阵, A 为任一 n 阶矩阵, x 为任一 n 维向量, 记 A^H, x^H 为其共轭转置。

定义 1.1 设 A 是 n 阶矩阵, 若存在非零向量 x , 数 λ 使得

$$Ax = \lambda Mx \quad (1.1)$$

成立, 称 λ 是 A 相对于 M 的广义特征值, 非零向量 x 称为属于 λ 的广义特征向量。

式(1.1)成立当且仅当关于 λ 的 n 次代数方程

$$|\lambda M - A| = 0 \quad (1.2)$$

成立。称式(1.2)是 A 相对于 M 的特征方程。

由定义不难得到, 对于给定的 M , 任一矩阵在复数域范围内恒有 n 个广义特征值(重数按重数计算)。

命题 1.1 A 相对于 M 的广义特征值等价于矩阵 $S = M_1^H A M_1$ 的特征值, 且与 M_1 的选取无关, 其中 M_1 是满足 $M_1^H M M_1 = E$ 的矩阵。

证 因为 M 是正定埃米特矩阵, 因此, 存在矩阵 M_1 , 使得 $M_1^H M M_1 = E$, 记 $S = M_1^H A M_1$, $u_1, u_2, \dots, u_n$ 是特征方程 $|\lambda E - S| = 0$ 的根, 即是矩阵 S 的特征根, 这样有 $|\lambda M_1^H M M_1 - M_1^H A M_1| = 0$, 化简得 $|\lambda M - A| = 0$ 。这说明 $u_1, u_2, \dots, u_n$ 是方程 $|\lambda M - A| = 0$ 的根, 即为 A 的相对于 M 的广义特征值。

反之, 设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是方程 $|\lambda M - A| = 0$ 的根, 则有 $|\lambda M_1^H M M_1 - M_1^H A M_1| = 0$, 即得 $|\lambda E - S| = 0$ 。这说明 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 S 的特征值。所以 S 的特征值与 A 的相对于 M 的广义特征值一致, 且从证明过程可知与 M_1 的选取无关。

定理 1.2 设 y 是 S 属于特征值 λ 的特征向量, 则 $M_1 y$ 是 A 属于广义特征值 λ 的广义特征向量; 反之, 若 y 是 A 属于广义特征值 λ 的广义特征向量, 则 $M_1^{-1} y$ 是 S 属于特征值 λ 的特征向量。

证 设 y 是属于 S 特征值 λ 的特征向量, 则 $Sy = \lambda y$, 即 $M_1^H A M_1 y = \lambda y$, 因此有 $M_1^H A M_1 y = \lambda M_1^{-1} M_1 y$, 亦即 $A M_1 y = \lambda (M_1^H)^{-1} M_1^{-1} M_1 y$ 。由于 $M_1^H M M_1 = E$, 故 $M = (M_1^H)^{-1} M_1^{-1}$, 因此, 有

$AM_1y = \lambda MM_1y$ 。这说明 M_1y 是 A 属于 λ 的广义特征向量。反之同理可证。

定义 1.2 称向量 x, y 是广义正交的, 若 $x^H My = 0$ 。

由定理 1.1 及定理 1.2 不难得到下述事实。

定理 1.3 (1) A 的属于不同广义特征值的广义特征向量线性无关。

(2) 设 A 为埃米特矩阵, 则 A 的广义特征值均为实数; 且有 n 个线性无关的广义特征向量, 属于不同广义特征值的广义特征向量是广义正交的。

(3) 设 A 为仅埃米特矩阵, 则 A 的广义特征值均为纯虚数; 且有 n 个线性无关的广义特征向量, 属于不同广义特征值的广义特征向量是广义正交的。

设 A 为埃米特矩阵(或反埃米特矩阵), 则 $M_1^H A M_1 = S$ 也是埃米特矩阵(反埃米特), 设 $y_1, y_2, \dots, y_n$ 是 S 的 n 个单位正交特征向量组, 即 $y_i^H y_j = \delta_{ij} (i, j = 1, 2, \dots, n)$ 。

由于 $M_1 y_i (i = 1, 2, \dots, n)$ 是 A 的广义特征向量, 且有 $(M_1^{-1} M_1 y_i)^H (M_1^{-1} M_1 y_j) = \delta_{ij}$

即 $(M_1 y_i)^H (M_1^{-1})^H M_1^{-1} M_1 y_j = \delta_{ij}$

因此 $(M_1 y_i)^H M M_1 y_j = \delta_{ij}$

所以 $M_1 y_1, M_1 y_2, \dots, M_1 y_n$ 是 A 的两两广义正交的广义特征向量。令 $x_i = M_1 y_i (i = 1, 2, \dots, n)$,

则 $x_i^H M x_j = \delta_{ij}, x_i^H A x_j = \lambda_j \delta_{ij}$

其中 λ_i 是对应特征向量 y_i 的 S 的特征值。

若记 $T = (x_1, x_2, \dots, x_n)$, 则有

$$T^H M T = E, T^H A T = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

2 广义酉相似

定义 2.1 矩阵 T 称为广义酉矩阵, 如果 T 满足 $T^H M T = E$ 。

定义 2.2 设 A, B 为 n 阶矩阵, 若存在广义酉矩阵 T , 满足 $T^H A T = B$, 称 A 与 B 广义酉相似(相对于 M)。

定理 2.1 任一矩阵 A 均广义酉相似于上三角矩阵, 且对角线元素为 A 的广义特征值。

证 由施斜尔(schur)定理存在酉矩阵 U , 使得 $U^H S U$ 为上三角矩阵。其中对角线上元素为 S 的特征值, 取 $T = M_1 U$ 即可。

定理 2.2 矩阵 A 广义酉相似于对角阵当且仅当 A 有 n 个广义正交的广义特征向量。

证 由定义即得。

这样埃米特矩阵, 仅埃米特矩阵均广义酉相似于对角阵。一般地有以下定理。

定理 2.3 矩阵 A 广义酉相似于对角阵当且仅当 A 满足 $A^H B^{-1} A = A B^{-1} A^H$ 。

证 由于 A 广义酉相似于对角阵等同于 S 酉相似于对角阵。因为 S 酉相似于对角阵当且仅当

$$S^H S = S S^H$$

将 $S = M_1^H A M_1$ 代入, 即得当且仅当有

$$A^H B^{-1} A = A B^{-1} A^H$$

定理 2.4 设 A, B 广义酉相似于对角阵(相对于 M), 则 A 与 B 的克罗内克(Kronecker)积 $A \otimes B$ 也广义酉相似于对角矩阵(相对于 $M \otimes M$)。

证 因为 A, B 广义酉相似于对角矩阵, 存在广义酉矩阵 T_1, T_2 , 使得

$$T_1^H A T_1 = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

$$T_2^H B T_2 = \text{diag}(u_1, u_2, \dots, u_n)$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_n; u_1, u_2, \dots, u_n$ 分别为 A, B 的广义特征值. 记 $T = T_1 \otimes T_2$, 则 $T^H(M \otimes M)T = T_1^H M T_1 \otimes T_2^H M T_2 = E \otimes E = E$, 且

$$T^H(A \otimes B)T = T_1^H A T_1 \otimes T_2^H B T_2$$

由于对角矩阵的克罗内克积仍为对角矩阵, 故 $A \otimes B$ 广义酉相似于对角矩阵, 且 $A \otimes B$ 的广义特征值为 $\lambda_i u_j (i, j = 1, 2, \dots, n)$.

3 广义特征值的瑞利表示

定义 3.1 设 A 为埃米特矩阵, $x \neq 0$ 为 n 维向量, 称实数

$$R(x) = \frac{x^H A x}{x^H M x}$$

为埃米特矩阵 A 的瑞利商.

瑞利商是研究广义特征值的一个十分有效的工具. 设 A 为埃米特矩阵, 将 A 的广义特征值按递减顺序排列

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

定理 3.1 $R(x)$ 具有下述性质: (1) $R(kx) = R(x)$, k 为任意实数; (2) $\lambda_n \leq R(x) \leq \lambda_1$; (3) $\min_{x \neq 0} R(x) = \lambda_n, \max_{x \neq 0} R(x) = \lambda_1$.

证 (1) 由定义即得

(2) 因为 A 是埃米特矩阵, 因此存在广义酉矩阵 T , 使得 $T^H A T = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. 令 $x = Ty$, 则

$$R(x) = \frac{y^H T^H A T y}{y^H T^H M T y} = \frac{y^H \Lambda y}{y^H y}$$

故对任意 y 均有

$$\lambda_n \leq \frac{y^H A y}{y^H y} = \frac{x^H A x}{x^H M x} \leq \lambda_1$$

(3) 取 $y = (1, 0, \dots, 0)$, $R(x)$ 达到极大值 λ_1 , 取 $y = (0, 0, \dots, 0, 1)$, $R(x)$ 达到极小值, 所以

$$\min_{x \neq 0} R(x) = \lambda_n, \quad \max_{x \neq 0} R(x) = \lambda_1$$

上述定理告诉我们, 对于瑞利 $R(x)$ 只要在单位球面 $x^H x = 1$ 上考虑即可.

定理 3.2 设 $x_1, x_2, \dots, x_n$ 是 A 的广义特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 所对应的广义特征向量, $L(x_1, x_2, \dots, x_{k-1})$ 表示由 $x_1, x_2, \dots, x_{k-1}$ 生成的子空间, 它的正交补子空间用 L_k 表示, 则

$$\lambda_k = \max_{x \in L_k} R(x)$$

证 因为 A 是埃米特矩阵, 由本文 1, 可设 $x_1, x_2, \dots, x_n$ 满足 $x_i^H M x_j = \delta_{ij}$, $x_i^H A x_j = \lambda_j \delta_{ij}$.

即取 $T = (x_1, x_2, \dots, x_n)$, 则

$$T^H M T = E, \quad T^H A T = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

对任意向量 x , 设 $x = c_1 x_1 + c_2 x_2 + \dots + c_n x_n = TC$, 其中 $C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$, 从而有

$$R(x) = \frac{(TC)^H A (TC)}{(TC)^H M (TC)} = \frac{C^H \wedge C}{C^H C} = \frac{\lambda_1 \bar{C}_1 C_1 + \dots + \lambda_n \bar{C}_n C_n}{\bar{C}_1 C_1 + \dots + \bar{C}_n C_n}$$

令 $a_i = \frac{\bar{C}_i C_i}{\bar{C}_1 C_1 + \dots + \bar{C}_n C_n} \geq 0$, 则 $\sum_{i=1}^n a_i = 1$.

当 $k=1$ 时, L_1 表示整个 n 维空间, 此时结论为定理 3.1. 当 $k=2$ 时, $x \in L_2$, 这时 $c_1=0$, 从而

$$R(x) = a_2 \lambda_2 + \dots + a_n \lambda_n \leq \lambda_2$$

取 $a_2=1, a_3=a_4=\dots=a_n=0$, 可以达到 λ_2 , 所以 $\max_{x \in L_2} R(x) = \lambda_2$. 同理可证当 $k=3, 4, \dots, n$.

关于广义特征值的其它一些表示式, 在此不再一一列举, 下面给出一个应用例子.

例 3.1 设 A, B 是埃米特矩阵, 试证

$$\lambda_{\min}(A+B) \geq \lambda_{\min}(A) + \lambda_{\min}(B)$$

其中 $\lambda_{\min}(A)$ 表示 A 相对于 M 的广义特征值中最小者.

证 因为

$$\begin{aligned} \lambda_{\min}(A+B) &= \min_{x \neq 0} \frac{x^H (A+B)x}{x^H Mx} = \min_{x \neq 0} \left(\frac{x^H Ax}{x^H Mx} + \frac{x^H Bx}{x^H Mx} \right) \\ &\geq \min_{x \neq 0} \frac{x^H Ax}{x^H Mx} + \min_{x \neq 0} \frac{x^H Bx}{x^H Mx} \\ &= \lambda_{\min}(A) + \lambda_{\min}(B) \end{aligned}$$

最后介绍一种求实对称矩阵的广义特征值及广义特征向量的初等变换方法.

设 A 为实对称矩阵, 由本文 1 不难得到下述事实: 存在实矩阵 T , 使得

$$T' M T = E, \quad T' A T = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

这样 T 的第 i 个列向量就是广义特征值 λ_i 的特征向量. 由此构造矩阵

$$\begin{pmatrix} B & A \\ E & 0 \end{pmatrix}$$

对上述矩阵中的 B 进行一次合同变换, 对 A 亦施以同样的合同变换, 但对 E 只施行初等列变换, 最后化为形如下述矩阵

$$\begin{pmatrix} E & \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \\ T & 0 \end{pmatrix}$$

则 T 的第 i 个列向量就是属于广义特征值 λ_i 的广义特征向量.

参考文献

- 1 P. Lancaster, M. Tismenetsky: The Theory of Matrices, New York 1984
- 2 Chris Rorrs, Howard Anton: Application of Linear Algebra, New York 1979
- 3 王朝瑞, 史荣昌. 矩阵分析. 北京理工大学出版社. 1989.