第 30 卷 第*期 Vol.30， No.*
2014 年 12 月 ***. 2014

关于一道线性代数试题的思考

李厚彪，李红，蒲和平，王转德
（电子科技大学 数学科学学院，四川 成都 611731）

[摘 要] 从一道线性代数试题出发，对其剖析，并深入探讨了矩阵零化多项式与所对应矩阵之间的内在联系，推广了已知的相关结果．最后，探讨了一般情形下求解此类问题的方法及相关结论．

[关键词] 矩阵零化多项式；Jordan标准形；特征向量

[中图分类号] O151 [文献标识码] A [文章编号] 1672-1454（2015）01-

1 引 言

试题：设三阶方阵 $\boldsymbol{A}$ 的特征值-1,1 对应的特征向量分别为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$，向量 $\boldsymbol{\alpha}_3$ 满足 $\boldsymbol{A\alpha}_3=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$. (1) 证明：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关； (2) 设 $\boldsymbol{P}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]$，求 $\boldsymbol{P}^{-1}\boldsymbol{AP}$.

对于第一问，很多同学可用反正法或根据线性无关的定义，给出相关证明；但对于第二问，很多同学，给出了如下解答：

解答：因 $\boldsymbol{A\alpha}_3=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$，易知 $\boldsymbol{\alpha}_2=\boldsymbol{A\alpha}_3-\boldsymbol{\alpha}_3=(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{\alpha}_3$，由于 $\boldsymbol{\alpha}_2$ 所对应的特征值是 1，从而 $(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{\alpha}_3=1\boldsymbol{\alpha}_2=\boldsymbol{A\alpha}_2=\boldsymbol{A}(\boldsymbol{A\alpha}_3-\boldsymbol{\alpha}_3)$，即 $(\boldsymbol{A}^2-2\boldsymbol{A})\boldsymbol{\alpha}_3=-\boldsymbol{\alpha}_3$，因此 $\boldsymbol{\alpha}_3$ 为 $\boldsymbol{A}^2-2\boldsymbol{A}$ 的特征向量，亦为矩阵 $\boldsymbol{A}$ 的特征向量；令 $\boldsymbol{A\alpha}_3=\lambda\boldsymbol{\alpha}_3$，将其带入 $(\boldsymbol{A}^2-2\boldsymbol{A})\boldsymbol{\alpha}_3=-\boldsymbol{\alpha}_3$ 得 $\lambda^2-2\lambda=-1$，从而 $\lambda=1$ 为矩阵 $\boldsymbol{A}$ 的二重特征值.

由上可知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 皆为 $\boldsymbol{A}$ 的特征向量，且对应的特征值分别为-1,1,1；又由(1)可知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关，从而 $\boldsymbol{P}^{-1}\boldsymbol{AP}=diag(-1,1,1)$. 证毕.

显然上述解答是错误的，下面我们首先剖析一下其错误的原因，并对其进一步推广.

2 问题分析

首先，可能大多数学生比较熟悉的是下列结果：

引理 2.1[1] 设 n 阶矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A\alpha}=\lambda_i\boldsymbol{\alpha}(\boldsymbol{\alpha}\neq 0)$，即特征值为 λ_i（$i=1,2,\dots,n$），对应的特征向量为 $\boldsymbol{\alpha}$，则矩阵多项式

$$f(\boldsymbol{A})=a_n\boldsymbol{A}^n+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{I}$$

的特征值为 $f(\lambda_i)$（$i=1,2,\dots,n$），且 $\boldsymbol{\alpha}$ 亦为 $f(\boldsymbol{A})$ 的特征向量.

关于这个引理的证明，很多教科书[1,2]上都有：由 $\boldsymbol{A\alpha}=\lambda_i\boldsymbol{\alpha}(\boldsymbol{\alpha}\neq 0)$，显然

[收稿日期] 2015-1-15；[修改日期] 2015-06-16；
[基金项目] 国家自然科学基金资助项目(11101071)；电子科技大学教学研究项目(2013XJYSL026,Y03003023901002010).

$$
\begin{aligned}
f(A)\alpha &= a_n A^n\alpha + \cdots + a_1 A\alpha + a_0 I\alpha \\
&= a_n\lambda_i^n\alpha + \cdots + a_1\lambda_i\alpha + a_0\alpha \\
&= (a_n\lambda_i^n + \cdots + a_1\lambda_i + a_0)\alpha \\
&= f(\lambda_i)\alpha,
\end{aligned}
$$

这表明 $f(\lambda_i)$ 是 $f(A)$ 的特征值，α 是 $f(A)$ 的特征向量. 下面让我们通过一个例子，理解一下该命题：

例1. 实对称矩阵 A 满足 $A^3+A^2+A=3I$，则 $A=$ ____.

解: 设矩阵 A 的特征值为 λ，则有 $\lambda^3+\lambda^2+\lambda=3$，即 $(\lambda-1)(\lambda^2+2\lambda+3)=0$. 由于实对称矩阵的特征值是实数，故 $\lambda^2+2\lambda+3=(\lambda+1)^2+2>0$，由此可得 A 只有唯一的三重特征值 1，因此存在可逆矩阵 P，使得 $P^{-1}AP=I$，即 $A=PP^{-1}=I$.

因此，由上面的引理 2.1，对于矩阵的零化多项式 $f(A)=O$ 来说，矩阵 A 的特征值一定是方程 $f(\lambda)$ 的根，A 的特征向量一定是 $f(A)$ 的特征向量. 但很多同学往往忽略了，上述逆命题是不一定成立的。例如：

例2. 设 $A=I$,显然 $f(A)=A^2-A=O$，而 $\lambda=0$ 是方程 $f(\lambda)=\lambda^2-\lambda=0$ 的根，但不是 A 的特征值.

例 3. $A=\begin{pmatrix}0&1\\1&0\end{pmatrix}$, $f(A)=A^2=\begin{pmatrix}1&0\\0&1\end{pmatrix}$, 显然任意二维非零向量皆是 $f(A)$ 的特征向量，但却不一定是 A 的特征向量，如 $A\begin{pmatrix}2\\3\end{pmatrix}=\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}2\\3\end{pmatrix}=\begin{pmatrix}3\\2\end{pmatrix}\neq k\begin{pmatrix}2\\3\end{pmatrix}$.

但为何出现 A 满足矩阵方程 $f(A)=O$，而 $f(\lambda)=0$ 的根又不是 A 的特征值的情况呢？让我们看下列两种情形：

(1) 若 A 满足矩阵方程 $g(A)=O$，则对 A 的任意多项式 $h(A)$，均有 $g(A)h(A)=O$；

(2) 由于两个非零矩阵的乘积也有可能为零矩阵，即 $g(A)\neq O$，$h(A)\neq 0$，但 $g(A)h(A)=O$.

因此，我们不能仅凭方程 $f(\lambda)=0$ 来确定 A 的特征值，除非 $f(\lambda)$ 就是 A 的特征多项式.

对前面的问题，事实上，我们很容易证明，α_3 并不是矩阵 A 的特征向量，即前面的解

答是错误的. 因假若 $A\alpha_3=\lambda\alpha_3$， 则由 $(A^2-2A)\alpha_3=-\alpha_3$ 得 $\lambda=1$ (二重特征值),即 $A\alpha_3=\alpha_3$，这显然与已知条件 $A\alpha_3=\alpha_2+\alpha_3(\alpha_2\neq 0)$ 矛盾.

一般说来，我们也很容易证明，矩阵多项式 $f(A)=O$ 的特征值和特征向量与矩阵 A 有如下关系：

定理 2.1. 若 A 的特征值 λ_0 是 $f(\lambda)=a_n\lambda^n+\cdots+a_1\lambda+a_0$ 的一个根，设 $f(\lambda)=\left(\lambda-\lambda_0\right)q\left(\lambda\right)$，则对任意满足 $f(A)u=0$ 的非零向量 u，若 $q(A)u\neq 0$，则 $q(A)u$ 是 A 的属于特征值 λ_0 的一个特征向量。

下面给出该问题的一种正确解法：

$$\begin{aligned}
P^{-1}AP&=[\alpha_1,\alpha_2,\alpha_3]^{-1}A[\alpha_1,\alpha_2,\alpha_3]\\
&=[\alpha_1,\alpha_2,\alpha_3]^{-1}[A\alpha_1,A\alpha_2,A\alpha_3]\\
&=[\alpha_1,\alpha_2,\alpha_3]^{-1}[-\alpha_1,\alpha_2,\alpha_2+\alpha_3]\\
&=[\alpha_1,\alpha_2,\alpha_3]^{-1}[\alpha_1,\alpha_2,\alpha_3]\begin{bmatrix}-1&0&0\\0&1&1\\0&0&1\end{bmatrix}\\
&=\begin{bmatrix}-1&0&0\\0&1&1\\0&0&1\end{bmatrix}.
\end{aligned}$$

3 关于该问题的进一步讨论

对于该问题，若学过矩阵的 Jordan 标准形的同学来说，可能是一个容易的问题。但对于 Jordan 标准形的引入和证明，国内目前有很多不同的版本，与此题关系比较密切的几个，可能平时并不引人注意，特简述如下，有兴趣的同学，可去进一步查找资料(如[2-5])：

定义 3.1[2,3] 设 A 是 n 阶矩阵，若存在一个数 λ，以及非零的 n 维列向量 α，使得

$$(A-\lambda I)^{k-1}\alpha\neq 0,\ (A-\lambda I)^k\alpha=0$$

则称 α 为矩阵 A 的属于特征值 λ 的 k 次**广义特征向量**. 特别当 $k=1$ 时，则 α 即为通常的特征向量.

定理 3.2[2,3] (**若尔当基定理**) 设 α_k 为矩阵 A 的属于特征值 λ 的 k 次广义特征向量, 定义

$$\alpha_i=(A-\lambda I)\alpha_{i+1},\qquad (i=k-1,k-2,\ldots,1)$$

则称线性无关的向量列 $\mathbf{T}_\lambda=\{\alpha_k,\alpha_{k-1},\cdots,\alpha_1\}$ 为 α_k 生成的**若尔当链**；且不同的特征值所对应的若尔当链线性无关，其并集 $\mathbf{T}=(T_{\lambda_1},T_{\lambda_2},\cdots,T_{\lambda_s})$ 形成矩阵 A 的**若尔当基**，并且满足

$$T^{-1}AT=\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix},\ \text{(Jordan 标准形)}$$

其中 J_i 是每个若尔当链所对应的若尔当形矩阵.

由上面的这些基本知识，对前面的问题，既然 $(A-I)\alpha_3=\alpha_2\neq 0,\ (A-\lambda I)^2\alpha_3=0$，我们不难得到，$\alpha_3$ 为矩阵 A 的属于特征值 1 的 2 次广义特征向量；又 α_1 是其特征值-1 所对应的特征向量(可看做 1 次广义特征向量)，因此由上面的定理 3.2，直接得到

$$P^{-1}AP=\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1}$$

另外，关于 Jordan 标准形中各若尔当矩阵的阶数，也可由下面的定理进行判断：

定理 3.3[5] 设 λ_0 为 n 阶方阵 A 的特征值，记 $B=A-\lambda_0 I$，则对任意正整数 k，A 的若尔当标准形中，主对角元为 λ_0，且阶数为 k 的 Jordan 块数等于：

$$rank(B^{k-1})-2rank(B^k)+rank(B^{k+1}).$$

对于此题中 $B=A-I$，由上面(1)式易知

$$P^{-1}BP=P^{-1}AP-I=\begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\Rightarrow rank(\mathrm{B})=2, \tag{2}$$

$$P^{-1}B^2P=(P^{-1}BP)^2=\begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\Rightarrow rank(\mathrm{B}^2)=1, \tag{3}$$

$$P^{-1}B^3P=(P^{-1}BP)^3=\begin{bmatrix} 8 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\Rightarrow rank(\mathrm{B}^3)=1, \tag{4}$$

因此，根据定理 3.3，矩阵 A 的 Jordan 标准形中，特征值 1 所对应的若尔当矩阵的阶数只能是 2. 另外，由(2)我们也可以看到：一个不可逆矩阵其秩的大小，并不等于非零特征值的个数(但对实对称矩阵来说，成立)，而是等于矩阵的阶数减去零特征值的几何重数(即 Jordan 标准形中以 0 为特征值的 Jordan 块的个数). 这一点也容易被初学者忽略，在此提醒一下.

[参 考 文 献]

[1] 黄廷祝,成孝予, 线性代数与空间解析几何(第三版)[M]. 高等教育出版社, 2008.
[2] 邱森,朱林生.高等代数探究性课题精编[M].武汉大学出版社，2012.
[3] Meyer C.D. Matrix Analysis and Applied Linear Algebra[M].Society for Industrial and Applied Mathematics.Philadelphia,PA,2002.

[4] Hall J.I. Another elementary approach to the Jordan form [J]. Amer. Math. Monthly,1991(98): 336-340.
[5] H.D. Ikramov. Linear Algebra: Problems book. [M]. Moscow, 1975.

A Discussion about a Question of Linear Algebra

LI Hou-biao, *LI Hong*, *PU He-ping*, *WANG Zhuan-de*
(School of Mathematics Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: Based on a linear algebra question, this paper discussed deeply the intrinsic link between the matrix zero polynomials and the corresponding matrix, which promotes the known results. In addition, the method and related conclusions for solving such problems under general cases are also discussed.

Key words: matrix zero polynomials; Jordan form; eigenvectors

通信作者：李厚彪，电子科技大学，数学科学学院
地址：成都市高新西区西源大道 2006 号(电子科技大学清水河校区)
邮编：611731
Tel：13194884912，Email：lihoubiao0189@163.com.