应用数学和力学, 第 25 卷 第 3 期(2004 年 3 月)
Applied Mathematics and Mechanics

应用数学和力学编委会编
重 庆 出 版 社 出 版

文章编号: 1000_0887(2004) 03_0233_06

用奇异值分解方法计算具有重特征值矩阵的特征矢量*

迟 彬, 叶庆凯

(北京大学 力学与工程科学系, 北京 100871)

(我刊编委叶庆凯来稿)

摘要: 若当(Jordan) 形是矩阵在相似条件下的一个标准形, 在代数理论及其工程应用中都具有十分重要的意义· 针对具有重特征值的矩阵, 提出了一种运用奇异值分解方法计算它的特征矢量及若当形的算法· 大量数值例子的计算结果表明, 该算法在求解具有重特征值的矩阵的特征矢量及若当形上效果良好, 优于商用软件 MATLAB 和 MATHEMATICA·

关 键 词: 重特征值; 特征矢量; 特征矢量链; 若当形
中图分类号: O151.21 文献标识码: A

引 言

在代数理论及其工程应用中, 矩阵的特征矢量及若当形越来越受到重视· 在振动系统的模态分析和控制系统的设计问题中, 更需要考虑如何计算具有重特征值的矩阵的特征矢量· 在不存在重特征值的情况下, 已有多种方法求矩阵的特征值及其相应的特征矢量, 由此可以求得矩阵的若当形, 并已研制出了广泛使用的商品软件, 如: EISPACK, MATLAB[1] 和 MATHEMATICA[2] 等· 在 MATLAB 的最新版本中, 也考虑了重特征值的情况· 但当阶数超过 5 且有重特征值时, 用 MATLAB 和 MATHEMATICA 软件计算的结果往往不尽如人意, 是误差较大的近似解· 造成上述误差的主要原因在于这些软件在计算特征矢量时采用的是反幂法[3], 原则上说不适用于重特征值的情况, 而用符号运算时, 矩阵元素又必须是整数或小的整数之比· 不适用于工程实际问题·

参考文献[4] 中给出了求矩阵重特征值的方法, 并取得了满意的结果· 在此基础上, 本文采用奇异值分解方法求矩阵的特征矢量及若当标准形· 算法的主要思想如下: 首先求矩阵的重特征值, 再运用奇异值分解方法确定这些重特征值的特征矢量及相应的若当块的个数· 然后利用若当块的特殊形式, 确定矩阵对应于每一个若当块的特征矢量链, 最后求出矩阵的特征矢量及若当形·

本文针对具有重特征值的矩阵, 提出了一种运用奇异值分解方法计算它的特征矢量及若

* 收稿日期: 2002_03_26; 修订日期: 2003_11_18
基金项目: 国家自然科学基金资助项目(69974003); 高等院校博士点基金资助项目(2001001011)
作者简介: 迟彬(1974 —), 女, 辽宁盖州人, 博士(联系人. Tel: 86_512_65321234_3575; Fax: 86_512_65321234_3555; E_mail: bin.chi@samsung.com) .

当形的算法• 数例表明, 本文提出的算法在求解具有重特征值的矩阵的特征矢量及若当形上效果良好•

1 算法的基本原理

由代数理论可知, 对于任意一个 n 阶方阵 $\boldsymbol{A}$, 如果 $\boldsymbol{A}$ 具有 n 个互不相同的特征值, 则 $\boldsymbol{A}$ 一定相似于一个对角矩阵, 即存在可逆矩阵 $\boldsymbol{W}$, 使得 $\boldsymbol{W}^{-1}\boldsymbol{A}\boldsymbol{W}$ 为对角形, 并且主对角线上元素为矩阵 $\boldsymbol{A}$ 的特征值; 而如果 $\boldsymbol{A}$ 的相异特征值的个数小于 n, 也就是矩阵 $\boldsymbol{A}$ 的特征多项式具有重根时, 则矩阵是否可以对角化需要视重特征值所对应的特征矢量空间的维数而定• 当特征矢量空间的维数之和等于方阵 $\boldsymbol{A}$ 的阶数时, $\boldsymbol{A}$ 仍可相似于一个对角矩阵, 主对角线上的元素除排列次序外是确定的, 它们是矩阵 $\boldsymbol{A}$ 的全部特征值(重根按重数计算)[5~6]•

当特征矢量空间的维数之和小于 $\boldsymbol{A}$ 的阶数时, $\boldsymbol{A}$ 不可对角化• 但是我们仍然可以找到一个可逆矩阵 $\boldsymbol{W}$, 使得 $\boldsymbol{W}^{-1}\boldsymbol{A}\boldsymbol{W}$ 具有如下的形式[6]:

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}_1 & & \boldsymbol{0} \\ & \boldsymbol{J}_2 & \\ & & \ddots \\ \boldsymbol{0} & & \boldsymbol{J}_t \end{bmatrix},$$

其中:

$$\boldsymbol{J}_i = \begin{bmatrix} \boldsymbol{J}_{i1} & & \boldsymbol{0} \\ & \boldsymbol{J}_{i2} & \\ & & \ddots \\ \boldsymbol{0} & & \boldsymbol{J}_{in_i} \end{bmatrix}, \quad \boldsymbol{J}_{ij} = \begin{bmatrix} s_i & 1 & & \boldsymbol{0} \\ & s_i & 1 & \\ & & \ddots & 1 \\ \boldsymbol{0} & & & s_i \end{bmatrix},$$

t 为矩阵 $\boldsymbol{A}$ 的相异特征值的个数, n_i 为对应于特征值 s_i 的特征矢量空间的维数, 方阵 $\boldsymbol{J}_{ij}$ $(i = 1, 2, \cdots, t;\ j = 1, 2, \cdots, n_i)$ 为对应于特征值 s_i 的若当块, 它的大小是 n_{ij} ($\sum_{j=1}^{n_i} n_{ij}$ 是特征值 s_i 的重数)• 称具有上述形式的矩阵 $\boldsymbol{J}$ 为矩阵 $\boldsymbol{A}$ 的若当标准形(Jordan), 这个若当形矩阵中除去若当块 $\boldsymbol{J}_{ij}$ 的排列次序外, 是唯一的•

要想求得矩阵 $\boldsymbol{A}$ 的若当标准形, 关键问题在于找到可逆矩阵 $\boldsymbol{W}$•

首先分析若当形标准形中的若当块 $\boldsymbol{J}_{ij}$• 若矩阵 $\boldsymbol{W}$ 存在, 则对于 $\boldsymbol{W}$ 中的相应于若当块 $\boldsymbol{J}_{ij}$ 的那部分列矢量, 下述等式成立:

$$\boldsymbol{A}[\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \cdots \quad \boldsymbol{v}_m] = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \cdots \quad \boldsymbol{v}_m]\boldsymbol{J}_{ij},$$

其中: $[\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \cdots \quad \boldsymbol{v}_m]$ 为 $\boldsymbol{W}$ 的部分列矢量组, $m = n_{ij}$• 因而有:

$$\boldsymbol{A}\boldsymbol{v}_1 = s_i\boldsymbol{v}_1, \ \boldsymbol{A}\boldsymbol{v}_j = \boldsymbol{v}_{j-1} + s_i\boldsymbol{v}_j \qquad (i = 1, 2, \cdots, t;\ j = 2, 3, \cdots, m)• \tag{1}$$

对上述公式进行转换, 得到:

$$(\boldsymbol{A} - s_i\boldsymbol{I})\boldsymbol{v}_1 = 0, \ (\boldsymbol{A} - s_i\boldsymbol{I})\boldsymbol{v}_j = \boldsymbol{v}_{j-1} \qquad (i = 1, 2, \cdots, t;\ j = 2, 3, \cdots, m)• \tag{2}$$

由此, 求矩阵 $\boldsymbol{W}$ 的问题被转化为求 $\boldsymbol{W}$ 中对应于各个若当块 $\boldsymbol{J}_{ij}$ 的列矢量组的问题•

在公式(2)中, $(\boldsymbol{A} - s_i\boldsymbol{I})\boldsymbol{v}_1 = 0$• $\boldsymbol{v}_1$ 为矩阵 $\boldsymbol{A}$ 对应于特征值 s_i 的特征矢量, $\boldsymbol{v}_1$ 的值可以通过奇异值分解得到, 关键在于 $\boldsymbol{v}_2, \boldsymbol{v}_3, \cdots, \boldsymbol{v}_m$ 的计算, 称满足公式(2)的 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ 为特征矢量 $\boldsymbol{v}_1$ 的特征矢量链•

显然, 可用如下方法计算特征矢量链:

1) 设有矩阵 $\boldsymbol{A}$, 它有 n 重特征值 s• 如果矩阵 $\boldsymbol{A}-s\boldsymbol{I}$ 有 n_1 个零奇异值, 则对应于特征值 s 共有 n_1 个若当块, 且相当于这些零奇异值的奇异矢量为对应于特征值 s 的特征矢量•

2) 设有矩阵 $\boldsymbol{A}$, s 是它的重特征值, $\boldsymbol{v}_1$ 是对应的特征矢量• 令

$$\boldsymbol{v}_i = (\boldsymbol{A}-s\boldsymbol{I})^{-1}\boldsymbol{v}_{i-1} \qquad (i=2,3,\cdots)\bullet$$

若 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ 线性无关, 而 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{m+1}$ 线性相关, 则对应于特征值 s 和它的特征矢量 $\boldsymbol{v}_1$ 的特征矢量链是 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$, 相应的若当块的大小为 $m\times m$•

2 算法的计算步骤

1) 给定要求特征矢量及若当形的 n 阶矩阵 $\boldsymbol{A}$•

2) 按文献[4]中的方法, 求矩阵 $\boldsymbol{A}$ 的相异特征值• 为此首先确定矩阵 $\boldsymbol{A}$ 的特征多项式 $f(s)=|s\boldsymbol{I}-\boldsymbol{A}|$, 并求它的导数多项式 $g(s)$, 再求出 $f(s)$ 和 $g(s)$ 的最大公因式 $d(s)$, 计算 $f(s)$ 除以 $d(s)$ 的商 $c(s)$• 用QR 迭代方法确定 $c(s)$ 的伴矩阵的特征值, 它们都是单特征值• 这些特征值即矩阵 $\boldsymbol{A}$ 的相异特征值, 共有 t 个, 记为 $s_1, s_2, \cdots, s_t$•

3) 对于每一个特征值 $s_i(i=1,2,\cdots,t)$, 判别它是单根还是重根• 为此计算 $d(s_i)$, 若它不等于零, 则 s_i 是单根, 否则 s_i 是重根•

4) 构成矩阵 $\boldsymbol{E}=\boldsymbol{A}-s_i\boldsymbol{I}$, 并对矩阵 $\boldsymbol{E}$ 进行奇异值分解•

5) 若 s_i 是单根, $\boldsymbol{E}$ 有一个零奇异值, 相应的奇异矢量是矩阵 $\boldsymbol{A}$ 对应于特征值 s_i 的特征矢量•

6) 若 s_i 是重根, 设 $\boldsymbol{E}$ 具有 n_i 个零奇值, 对应于这些零奇异值的奇异矢量分别为 $\boldsymbol{v}_{i1}, \boldsymbol{v}_{i2}, \cdots, \boldsymbol{v}_{in_i}$• 它们是矩阵 $\boldsymbol{A}$ 对应于特征值 s_i 的特征矢量, n_i 是该特征值所对应的特征矢量空间的维数• 对于每一个 $\boldsymbol{v}_{ij}(i=1,2,\cdots,t;\ j=1,2,\cdots,n_i)$, 令

$$\boldsymbol{v}_1 = \boldsymbol{v}_{ij},$$

$$\boldsymbol{v}_k = (\boldsymbol{A}-s_i\boldsymbol{I})^{-1}\boldsymbol{v}_{k-1} \qquad (k=2,3,\cdots)\bullet$$

逐步迭代的过程中得到矢量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ 线性无关, 而 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{m+1}$ 线性相关• 则 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ 为对应于特征值 s_i 及其特征矢量 $\boldsymbol{v}_{ij}$ 的特征矢量链•

7) 将所求得的全部特征矢量(包括特征矢量链)合并成一个矩阵 $\boldsymbol{W}$•

8) 计算 $\boldsymbol{J}=\boldsymbol{W}^{-1}\boldsymbol{A}\boldsymbol{W}$, 它即是矩阵 $\boldsymbol{A}$ 的若当形•

3 数 值 例 子

为了验证上述算法, 我们运用该算法计算了大量的实例, 并将计算结果与 MATLAB 和 MATHEMATICA 的结果进行了比较•

考虑如下形式的矩阵

$$\boldsymbol{A}(s)=\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ s^5 & -5\times s^4 & 10\times s^3 & -10\times s^2 & 5\times s \end{bmatrix},$$

理论上, 它的五个特征值都是 s, 且只有一个若当块:

$$
\boldsymbol{J}(s)=\begin{bmatrix} s & 1 & 0 & 0 & 0 \\ 0 & s & 1 & 0 & 0 \\ 0 & 0 & s & 1 & 0 \\ 0 & 0 & 0 & s & 1 \\ 0 & 0 & 0 & 0 & s \end{bmatrix}\text{•}
$$

例 1

考虑矩阵 $\boldsymbol{A}$ (3.23)•

将数值3.23 代入到 $\boldsymbol{A}(s)$ 中,得到:

$$
\boldsymbol{A}(3.23)=\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 351.570\,649\,784\,3 & -544.227\,012\,05 & 336.982\,67 & -104.329 & 16.15 \end{bmatrix}\text{•}
$$

用文献[4]中的方法得到它的五个特征值都是3.23,因而其相异特征值只有一个3.23•

构成矩阵 $\boldsymbol{A}-3.23\boldsymbol{I}$• 计算得到矩阵只有一个零奇异值,表明它只有一个若当块• 零奇异值所对应的奇异矢量为相应于特征值3.23的特征矢量,用我们的程序可以求得它的特征矢量链并构成矩阵 $\boldsymbol{W}$• 可以用计算误差 $\varepsilon=\|\boldsymbol{W}^{-1}\boldsymbol{A}\boldsymbol{W}-\boldsymbol{J}(3.23)\|_2$ 的方法判别 $\boldsymbol{W}$ 的正确性• 对此例我们得到 $\varepsilon=1.467\,1\times10^{-12}$, 可见计算结果是正确的•

用MATLAB软件(调用符号运算包的Jordan函数)对矩阵 $\boldsymbol{A}(3.23)$ 进行计算,得到矩阵的近似若当标准形 $\boldsymbol{J}_T$,

$\boldsymbol{J}_T=$

$$
\begin{bmatrix} 3.226\,7 & 0 & 0 & 0 & 0 \\ 0 & 3.229\,0-0.003\,1\mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & 3.229\,0+0.003\,1\mathrm{i} & 0 & 0 \\ 0 & 0 & 0 & 3.232\,7-0.001\,9\mathrm{i} & 0 \\ 0 & 0 & 0 & 0 & 3.232\,7+0.001\,9\mathrm{i} \end{bmatrix}\text{•}
$$

计算 $\varepsilon=\|\boldsymbol{J}_T-\boldsymbol{J}(3.23)\|_2=1.002\,7$• 误差估计显示计算结果有较大的偏差•

同样用MATHEMATICA软件对矩阵 $\boldsymbol{A}(3.23)$ 进行计算,也得到了它的近似若当标准形 $\boldsymbol{J}_A$:

$\boldsymbol{J}_A=$

$$
\begin{bmatrix} 3.226\,92-0.002\,2\mathrm{i} & 0 & 0 & & 0 \\ 0 & 3.226\,92+0.002\,2\mathrm{i} & 0 & & 0 \\ 0 & 0 & 3.231\,17-0.003\,6\mathrm{i} & & 0 \\ 0 & 0 & & 3.231\,17+0.003\,6\mathrm{i} & 0 \\ 0 & 0 & & & 3.233\,81 \end{bmatrix}\text{•}
$$

误差估计值 $\varepsilon=\|\boldsymbol{J}_A-\boldsymbol{J}(3.23)\|_2=1.001\,8$, 表明计算结果是一个误差较大的近似解•

我们对具有上述特点的矩阵进行了大量的计算,结果显示出我们的算法取得了令人满意的结果• 现将该算法与MATLAB和MATHEMATICA的计算结果列于表1中•

表中"Jordan"代表计算的结果为正确的若当标准形,其相应的误差值 $\varepsilon<10^{-11}$; "对角形"代表计算结果若当形为对角矩阵; "NaN"代表"Not_a_Number",表明计算过程中遇到了数值问题(如: 0/0, 0×inf 等)•

表 1 计算矩阵 $\boldsymbol{A}(s)$ 的特征矢量与若当形

s 值	奇异值分解算法	MATLAB	MATHEMATICA
0.1; 0.3; 0.4; 0.6; 0.9	Jordan	Jordan	对角形
0.28; 1.1; 2.1; 2.17; 2.2; 3.23; 4.2; 4.7; 5.1; 6.9; 7.4; 8.2	Jordan	对角形	对角形
0.334; 0.57; 1.28; 2.7; 3.3; 3.6	Jordan	NaN	对角形

从表 1 中看到, 奇异值分解算法得到了具有重特征值的矩阵的特征矢量(特征矢量链)及若当标准形•

例 2

为了进一步验证算法, 我们扩大矩阵的阶数到 10, 形式如下:

$$\begin{bmatrix} \boldsymbol{A}(s) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}(s) \end{bmatrix},$$

其中, 每一个 $\boldsymbol{A}(s)$ 为形如例 1 的 5 阶矩阵, 对表 1 中的所有 s 值进行计算, 结果显示该矩阵的特征值是十重, 有两个阶数为 5 的若当块• 计算结果仍然是正确的•

计算结果分析:

对于具有重特征值的矩阵, 特别当重特征值所对应的若当块 $\boldsymbol{J}_{ij}$ 的阶数 n_{ij} 较大时, 利用我们提出的奇异值分解算法在计算具有重特征值的矩阵的特征矢量及若当标准形上显示出了明显的优势, 计算结果是令人满意的, 优于软件 MATLAB 和 MATHEMATICA•

4 结 论

本文在参考文献[4]的基础上, 不但求得了矩阵的重特征值, 而且计算出了对应于这些特征值的重特征矢量(特征矢量链), 进而计算得到了矩阵的若当标准形• 大量的数例表明, 本文所提出的算法在求解具有重特征值的矩阵的特征矢量及若当形上效果良好•

本文提出的方法也可用于无重特征值的情况, 由于奇异值分解采用的是正交变换, 原则上讲, 其数值稳定性可能比 EISPACK 中采用的反幂法更好一些•

本文的工作是初步的, 当矩阵元素的大小相差很大, 即其条件很差时, 直接用本文的方法也不能得到满意的结果, 进行适当的标定可能是有效的, 这方面的工作还在进行中•

[参 考 文 献]

[1] The Math Works Inc. MATLAB user's Guide[M]. Natick, Mass, U S Inc, 1992.
[2] 沈凤贤, 丁英仁, 赵文晖. Mathematica 手册[M]. 北京: 海洋出版社, 1992.
[3] 叶庆凯. 控制系统计算机辅助设计[M]. 北京: 北京大学出版社, 1990.
[4] 叶庆凯. 矩阵重特征值的一种计算方法[J]. 控制理论与应用, 1998, **15**(1): 118—120.
[5] 韩京清, 何关钰, 许可康. 线性系统理论代数基础[M]. 沈阳: 辽宁科学技术出版社, 1989.
[6] 北京大学数学力学系几何与代数科研室代数小组. 高等代数[M]. 北京: 人民教育出版社, 1978.

Computing the Eigenvectors of a Matrix With Multiplex Eigenvalues by SVD Method

CHI Bin, YE Qing_kai

(Department of Mechanics and Engineering Science, Peking University, Beijing 100871, P.R. China)

Abstract: Every matrix is similar to a matrix in Jordan canonical form, which has very important sense in the theory of linear algebra and its engineering application. For a matrix with multiplex eigenvalues, an algorithm based on the singular value decomposition (SVD) for computing its eigenvectors and Jordan canonical form was proposed. Numerical simulation shows that this algorithm has good effect in computing the eigenvectors and its Jordan canonical form of a matrix with multiplex eigenvalues. It is superior to MATLAB and MATHEMATICA.

Key words: multiplex eigenvalue; eigenvector; eigenvector chain; Jordan canonical form