关于 Rayleigh 商迭代法的收敛性

唐永才

(荆门职业技术学院 经管系,湖北 荆门 448000)

[摘 要] 文[1]讨论了 Rayleigh 商迭代法的收敛性,但在给出的说明中,对酉矩阵 Q 的形式提出了一些不适当的要求,额外附加了若干限制. 本文改进了文[1]中关于 商迭代法二次收敛性的证明.

[关键词] 商;商迭代法;二次收敛性.

[中图分类号] O241.6 [文献标识码] A [文章编号] 1008-4657(2005)06-0069-04

本文采用文[1]中的标准记号,以下无特殊说明,我们的讨论在复空间 C^n 中进行. 文中的范数取 2-范数 $\|\bullet\|_2$

众所周知,逆幂法的一个重要变形是 Rayleigh 商迭代法,其迭代格式是

$$(\lambda I - A)x' = \sigma x$$

其中向量 x' 作为初始向量 x 的一次迭代的结果,σ 为一适当的比例因子,λ 为矩阵 A 的某一近似特征值. 假如没有 A 的特征值的一个好的近似值,但已经知道初始向量 x 是相当精确的近似向量,则 Rayleigh 商 $\lambda = \dfrac{x^H Ax}{x^H x}$ 是对应于近似特征向量 x 的特征值的一个好的近似值. 利用这个 λ 值,解方程组 $(\lambda I - A)x' = \sigma x$,就得到一个新的特征向量. 由新的特征向量再计算一个新的 Rayleigh 商,如此循环反复,这便是逆幂法的 Rayleigh 商迭代过程.

Rayleigh 商迭代过程是否收敛,收敛究竟有多快,这是计算方法研究的一个重要课题. *Ostrowski,A* 在文[3]中曾有相当详尽的分析讨论. 不难看到下面给出的对 Rayleigh 商迭代法二次收敛性的证明是对文[1]中证明的较好的改进.

1 预备知识

定义 1.1 设序列 $\langle x^{(k)} \rangle$ 是一个 n 维向量序列,收敛于 $x \in C^n$,若存在与迭代次数 k 无关的数 $0 < \beta < \infty$ 和 $\alpha \geq 1$,使 k 从某个 $k_0 > 0$ 开始总有

$$\|x^{(k+1)} - x\|_2 \leq \beta \|x^{(k)} - x\|_2^{\alpha}$$

成立,则称 $\langle x^{(k)} \rangle$ 收敛的阶(次)为 α,或称序列 $\langle x^{(k)} \rangle \alpha$ 阶(次)收敛.

当 $\alpha = 2$ 时,称为二阶(次)收敛.

当 $1 < \alpha < 2$ 时,称为超线性收敛.

当 $\alpha = 1$,且 $0 < \beta < 1$ 时,称线性收敛或一阶(次)收敛.

定义 1.2 设 $A \in C^{n\times n}, x \in C^n$ 是 A 的对应于近似特征值 μ 的一个近似特征向量,则称

$$r = Ax - \mu x \tag{1}$$

为残差向量.

G. W. 斯图尔特在文[1]中证明了下面的定理.

定理 1.1 设 r 是由(1)所定义,那么当

[收稿日期] 2005-07-10
[作者简介] 唐永才(1946-),男,湖北荆门人,荆门职业技术学院副教授. 研究方向:数学规划. E-mail: tangyc6754@hotmail.com.

$$\mu = \frac{x^H A x}{x^H x} \tag{2}$$

时，$\| r \|_2$ 为极小值.

定义 1.3 由(2)式所定义的 μ 称为向量 x 的 Rayleigh 商.

定理 1.2 设 (x,U) 是酉阵,并假定 x 已规范化，$\| x \|_2 = 1$,U 的列正交于 x ,即 $U^H x = 0$,那么

$$\| g \|_2 = \| Ax - (x^H A x)x \|_2$$

其中

$$g = U^H A x$$

定理 1.3 设 x 是埃米特矩阵 A 的特征值 λ 对应的一个特征向量,如果

$$y = x + o(\varepsilon)$$

那么

$$\frac{y^H A y}{y^H y} = \lambda + 0(\varepsilon^2)$$

从这一定理看到,如果 y 是 A 的一个近似特征向量,那么 Rayleigh 商

$$\frac{y^H A y}{y^H y}$$

是 A 的一个近似特征值. 但是,当 A 是埃米特矩阵时,商

$$\frac{y^H A y}{y^H y}$$

作为一个近似特征值比 y 作为一个近似特征向量的精确度要高得多.

2 结果及证明

正因为上面的定理 1.3,Rayleigh 商迭代法具有较快的收敛速度,它是二次收敛的. 我们归纳为下面的定理 2.1.

定理 2.1 设 $A \in C^{n\times n}$,$x \in C^n$,矩阵 A 的近似特征向量由迭代格式

$$(\lambda I - A)x' = \sigma x \tag{3}$$

给出. 如果取 Rayleigh 商

$$\lambda = \frac{x^H A x}{x^H x}$$

则 Rayleigh 商迭代是二次收敛的.

其中 x' 是作为初始向量 x 进行一次迭代的结果，σ 为一适当的比例因子，λ 为 A 的某一近似特征值.

证明 设 R 为任意的酉阵. 对变量 x 进行以下变换,令

$$y = R^H x, y' = R^H x'$$

由(2)有

$$\lambda = \frac{y^H R^H A R y}{y^H y}$$

从(3)得到

$$(\lambda I - R^H A R)y' = \sigma y \tag{4}$$

解方程组(4)得 y' ,同样地由(2)有商

$$\lambda' = \frac{(y')^H (R^H A R) y'}{(y')^H y'} \tag{5}$$

为了方便,作如下变换:

假定 $\| x \|_2 = 1$,$R = (x,U)$ 为酉阵,则 $y = e_1$,且

$$R^H A R = \begin{pmatrix} \lambda & h^H \\ g & C \end{pmatrix} \tag{6}$$

利用(4),并适当选取比例因子σ,使得y'的第一个分量是一个单位,可以有

$$y' = (1,p^H)^H$$

这里$p = (\lambda I - C)^{-1}g$

由此可得y'是方程组

$$\begin{pmatrix} 0 & -h^H \\ -g & \lambda I - C \end{pmatrix}\begin{pmatrix} 1 \\ p \end{pmatrix} = \begin{pmatrix} \sigma \\ 0 \end{pmatrix}$$

的解.

其中

$$p = (\lambda I - C)^{-1}g \tag{7}$$

不失一般性,设利用$\dfrac{y'}{\|y'\|_2}$扩充成的酉矩阵为

$$Q = \begin{pmatrix} 1/\|y'\|_2 & u^H \\ p/\|y'\|_2 & V \end{pmatrix} \tag{8}$$

则由

$$Q^H(R^HAR)Q = \begin{pmatrix} \lambda' & (h')^H \\ g' & C' \end{pmatrix}$$

或者

$$\begin{pmatrix} 1/\|y'\|_2 & p^H/\|y'\|_2 \\ u & V^H \end{pmatrix}\begin{pmatrix} \lambda & h^H \\ g & C \end{pmatrix}\begin{pmatrix} 1/\|y'\|_2 & u^H \\ p/\|y'\|_2 & V \end{pmatrix} = \begin{pmatrix} \lambda' & (h')^H \\ g' & C' \end{pmatrix}$$

得到分块矩阵第二行第一列的子块:

$$\begin{aligned} g' &= (u,V^H)\begin{pmatrix} \lambda & h^H \\ g & C \end{pmatrix}(1/\|y'\|_2, p/\|y'\|_2)^T \\ &= \frac{\lambda u + V^Hg + (uh^H + V^HC)p}{\|y'\|_2} \\ &= \frac{\lambda u + V^H(\lambda I - C)p + uh^Hp + V^HCp}{\|y'\|_2} \\ &= \frac{\lambda[(1,p^H)\begin{pmatrix} u^H \\ V \end{pmatrix}]^H + uh^Hp}{\|y'\|_2} \\ &= \frac{uh^Hp}{\|y'\|_2} \end{aligned}$$

由定理1.1已知,作为R^HAR的特征向量y和y'的精确性的度量,分别是数

$$\rho = \frac{\|\lambda y - (R^HAR)y\|_2}{\|y\|_2}$$

和

$$\rho' = \frac{\|\lambda' y' - (R^HAR)y'\|_2}{\|y'\|_2}$$

由定理1.2,即有

$$\rho' = \|g'\|_2 \leqslant \frac{\|u\|_2\|h\|_2\|p\|_2}{\|y'\|_2} \leqslant \|u\|_2\|h\|_2\|p\|_2 \tag{9}$$

(这里$\|y'\|_2 = (1 + p^Hp)^{1/2} \geqslant 1$)

已知酉阵每行的2-范数$\|\cdot\|_2 = 1$,由(8)知Q的第一行的2-范数有

$$\frac{1}{\|y'\|_2^2} + \|u\|_2^2 = 1$$

于是

$$\| u \|_2^2 = 1 - \frac{1}{\| y' \|_2^2} = 1 - \frac{1}{1 + p^H p}$$

$$= \frac{p^H p}{1 + p^H p} \leqslant p^H p = \| p \|_2^2$$

即

$$\| u \|_2 \leqslant \| p \|_2 \tag{10}$$

将(10)代入(9),并注意到(7)及针对(6)应用定理1.2,即得

$$\rho' \leqslant \| h \|_2 \| p \|_2^2 \leqslant \| h \|_2 \| (\lambda I - C)^{-1} \|_2^2 \| g \|_2^2 = k_1 \rho^2$$

其中 $k_1 = \| h \|_2 \| (\lambda I - C)^{-1} \|_2^2$ 为一有界量.

特别指出,当 A 为埃米特矩阵时,在上面的推导中,从(6)式就有 $g = h$,因此又有

$$\rho' \leqslant \| h \|_2 \| (\lambda I - C)^{-1} \|_2^2 \| g \|_2^2 \leqslant k_2 \rho^3$$

其中 $k_2 = \| (\lambda I - C)^{-1} \|_2^2$ 为一有界量.

鉴于此,对于埃米特矩阵来说,Rayleigh 商迭代还是立方收敛的.

[参考文献]

[1] G. W. 斯图尔特著. 王国荣,黄丽萍,许锦龙,等译. 矩阵计算引论[M]. 上海:上海科学出版社,1980.

[2] 蒋尔雄,高坤敏,吴景琨. 线性代数[M]. 北京:人民教育出版社,1978.

[3] Ostrowski, A. On the convergence of the Rayleigh quotient iteration for the compuration of characterstic roots and vectors [J]. Ⅰ, Ⅱ, Ⅲ, Ⅳ, Ⅴ, Ⅵ, Arch Rational Mech . Anal. ,1958 ~ 1959.

On the Convergence of the Rayleigh Quotient Iteratire

TANG Yong - cai

(Jingmen Technical College, Jingmen, Hubei, 448000, China)

Abstract: The paper discusses the convergence of Rayleigh quotient and improves the confirmation of the features through eliminating some of the over demanding requirements and conditions.

Key words: Rayleigh quotient; Rayleigh quotient iteration; twice convergence

学术论文标题拟制的基本要求

标题是学术论文的有机组成部分,它是否规范,会对论文的质量带来很大的影响,因而必须引起我们的高度重视。

期刊文章的标题一般包括篇名(大标题)和文内标题(小标题)两种。篇名的结构形式又大体分为唯一标题型篇名和正题与副题组合而成的篇名。而文内标题则有一定的层次,可逐级分为标题、段首标题(提示标题)、细目标题三个层次;或分为一级、二级、三级……标题。论文标题的拟制须符合以下基本要求。

第一,准确妥贴。所谓准确,主要是指标题要能概括文意,达到文题相符。标题在概括文章的内容时不能失之过宽或过窄。所谓妥贴,主要体现为标题在反映论文的内容时要有层次性,即论文本身研究到什么范围、深入到何种程度,标题就反映到何种层次。

第二,新颖多样。从本质上讲,学术论文是作者科学研究有所发明、有所创造的成果的文字纪录。学术论文内容的创新性必然决定其标题的新颖性、多样性。也只有标题新颖、独特的论文才能引人注目。同时标题新颖性还要求题式多样,而题式多样又要求正确选择标题词语。

第三,简洁明了。所谓简洁,就是指语言的简明、洁净、精炼,要"文约而事丰"、言短而旨远。一个好的论文标题都要尽量达到直截了当、一语破的、一目了然。文章标题字数应控制在12字以内为宜。

另外,学术论文的篇名或文内标题都应避免使用不常见的缩略语,首字母缩写字、字符、代号、公式,应少用或不用冷僻字。文内标题层次不宜过多,一般3层,最多不超过5层。