中山大学学报 (自然科学版)
第 37卷　第 4期　ACTA SCIENTIARUM NATURALIUM　Vol. 37　No. 4
1998年　7月　UNIVERSITATIS SUNYATSENI　Jul.　1998

奇异值估计的 Brauer型定理*

李赛健　黎罗罗
(中山大学科学计算与计算机应用系，广州 510275)

摘　要　设 $A=(a_{ij})$为 n阶复矩阵. 记 $s_i=\max\left\{\sum_{j\neq i}|a_{ij}|, \sum_{j\neq i}|a_{ji}|\right\}$, $a_i=|a_{ii}|$ $(i=1,2,\cdots,n)$. 证明了 A的奇异值均属于 Brauer型并集 $\bigcup_{i\neq j}\left\{z\geqslant 0 \mid |z-a_i||z-a_j|\leqslant s_is_j\right\}$, 并给出了该并集的显式表达及数值例子.

关键词　奇异值, 包含区域, 条件数
分类号　O 151.21

奇异值估计在数值分析中有重要意义. 例如, 谱条件数 $k(A)=e_{\max}(A)$反映了方程组 $Ax=b$的病态程度, 并可借助对 $e_{\max}$的上界及 $e_{\min}$的下界的估计, 获得对谱条件数的上界估计. 关于特征值的估计, 有著名的 Gerschgorin圆盘定理, 相应地, 文 [3] 证明了 Gerschgorin型奇异值估计定理. 关于特征值估计, 还有更细微的 Brauer定理[1].

设 $A=(a_{ij})$是 n阶复方阵, 记 $r_i=\sum_{j\neq i}|a_{ij}|$.

引理 1　A的所有特征值均属于复平面上 $n(n-1)/2$个卵形区域的并集

$$\bigcup_{i\neq j}\left\{z\in \mathbf{C} \mid |z-a_{ii}||z-a_{jj}|\leqslant r_ir_j\right\} \tag{1}$$

关于 Brauer定理的评述, 见文 [4]. 本文将导出相应的 Brauer型奇异值估计定理. 以下设　$a_i=|a_{ii}|$, $u_+=\max(u,0)$及 $c_i=\sum_{j\neq i}|a_{ji}|$, $s_i=\max(r_i,c_i)$.

定理 1　矩阵 A的任一奇异值都必属于下述 $n(n-1)/2$个区域 $L_{ij}(A)$的并集 $L(A)=\bigcup_{i\neq j}L_{ij}(A)$.

$$L_{ij}(A)=\left\{z\geqslant 0 \mid |z-a_i||z-a_j|\leqslant s_is_j\right\} \qquad i\neq j \tag{2}$$

证明　设 e是 A的奇异值, 若 e等于某个 a_{i_0}, 则 $e\in L_{i_0j}(A)$　$j\neq i_0$, 定理结论成立. 下面设 $e\neq a_i$ $(i=1,2,\cdots,n)$. 因为 e是 A的奇异值, 故存在 2个非零向量 $\boldsymbol{x}=(x_1,\cdots,x_n)^{\mathrm{T}}$, $\boldsymbol{y}=(y_1,\cdots,y_n)^{\mathrm{T}}$使得

$$e\boldsymbol{x}=A^*\boldsymbol{y}, \qquad e\boldsymbol{y}=A\boldsymbol{x} \tag{3}$$

记 $z_j=\max(|x_j|,|y_j|)$ $(j=1,2,\cdots,n)$, p是使 $z_p=\max_{1\leqslant j\leqslant n}(z_j)$的指标, q是使 $z_q=\max_{j\neq p}(z_j)$的指标. 显然 $z_p\neq 0$ (否则 $\boldsymbol{x}$, $\boldsymbol{y}$为零向量), 另外也必有 $z_q\neq 0$, 因若不然则有 $x_j=0$, $y_j=0$ $(j\neq p)$, 这时由 (3) 式知 $ex_p=\bar{a}_{pp}y_p$及 $ey_p=a_{pp}x_p$ 得 $e^2=|a_{pp}|^2$即 $e=a_p$, 与假设矛盾.

* 收稿日期: 1997-07-10　李赛健, 男, 25岁, 博士研究生, 现在美国阿利桑纳州立大学

不失一般性，设 $z_p = |y_p|$， $z_q = |y_q|$ (其它情形如 $z_p = |y_p|$， $z_q = |x_q|$，… 等时证法相同).由(3)式的第 p 式得:

$$^{e}x_p - \bar{a}_{pp} y_p = \sum_{j \neq p} \bar{a}_{jp} y_j, \qquad ^{e}y_p - a_{pp} x_p = \sum_{j \neq p} a_{pj} x_j \tag{4}$$

记 $Z= x_p / y_p$，由 $|Z| \leqslant 1$，则 (4) 式知

$$|^{e}Z - \bar{a}_{pp}| \, |y_p| \leqslant \sum_{j \neq p} |\bar{a}_{jp}| \, |y_j|, \qquad |^{e} - Z a_{pp}| \, |y_p| \leqslant \sum_{j \neq p} |a_{pj}| \, |x_j| \tag{5}$$

由 z_p，z_q 的定义知

$$|^{e}Z - \bar{a}_{pp}| \leqslant z_q \sum_{j \neq p} |\bar{a}_{jp}| / z_p = c_p z_q / z_p, \qquad |^{e} - Z a_{pp}| \leqslant z_q \sum_{j \neq p} |a_{pj}| / z_p = r_p z_q / z_p \tag{6}$$

若 $^{e} \geqslant a_p$，则 $|^{e} - a_p| \leqslant |^{e} - |Z| a_p| \leqslant |^{e} - Z a_{pp}| \leqslant r_p z_q / z_p$；否则 $|^{e} - a_p| \leqslant ||Z|^{e} - a_p| \leqslant |Z^{e} - \bar{a}_{pp}| \leqslant c_p z_q / z_p$，但无论如何均有

$$|^{e} - a_p| \leqslant s_p z_q / z_p \tag{7}$$

类似地，由 (3) 式的第 q 式有

$$|^{e} - a_q| \leqslant s_q z_p / z_q \tag{8}$$

综合(7),(8)式得 $|^{e} - a_p| \, |^{e} - a_q| \leqslant s_p s_q$，可见 ^{e}属于 $L_{pq}(\boldsymbol{A})$从而 $^{e} \in L(\boldsymbol{A})$.

下面将给出区域 $L(\boldsymbol{A})$的显式表示.

引理 2 设 $0 \leqslant a < b$，$c = (a + b)/2$，$d = (b - a)/2$，$g \geqslant 0$，则不等式: $|z - a| \, |z - b| \leqslant g$，$z \geqslant 0$的解集为

$$\left[(c - \sqrt{d^2 + g})_+, c - \sqrt{(d^2 - g)_+}\right] \cup \left[c + \sqrt{(d^2 - g)_+}, c + \sqrt{d^2 + g}\right] = \left[(a - (\sqrt{(d^2 + g)} - d))_+, a + \left(d - \sqrt{(d^2 - g)_+}\right)\right] \cup \left[b - \left(d - \sqrt{(d^2 - g)_+}\right), b + \left(\sqrt{d^2 + g} - d\right)\right]$$

证明 注意 $|z - a| \, |z - b| = |(z - c)^2 - d^2|$，由 $|z - a| \, |z - b| \leqslant g$ 导得 $d^2 \leqslant (z - c)^2 \leqslant d^2 + g$ 或 $d^2 - g \leqslant (z - c)^2 \leqslant d^2$. 用初等运算可得到引理 2结论. 证毕.

以下设

$d_{ij} = (a_j - a_i)/2$，$l_{i1} = \max\left\{\sqrt{d_{ij}^2 + s_i s_j} - d_{ij} \mid a_j \geqslant a_i\right\}$，$l_{i2} = \max\left\{d_{ki} - \sqrt{(d_{ki}^2 - s_k s_i)_+} \mid a_k \leqslant a_i\right\}$，$u_{i1} = \left\{d_{ij} - \sqrt{(d_{ij}^2 - s_i s_j)_+} \mid a_j \geqslant a_i\right\}$，$u_{i2} = \max\left\{\sqrt{d_{ki}^2 + s_k s_i} - d_{ki} \mid a_k \leqslant a_i\right\}$，$l_i = \max(l_{i1}, l_{i2})$，$u_i = \max(u_{i1}, u_{i2})$，$V_i(\boldsymbol{A}) = V_i = [(a_i - l_i)_+, a_i + u_i]$.

定理 2 $\boldsymbol{A}$的任一奇异值属于区间并集 $\bigcup_{i=1}^{n} V_i$；而且，若 $\tilde{V} = \bigcup_{j=1}^{k} V_{i_j}$，为一连通的区间，$\tilde{V}$ 与其它 V_t ($t \neq i_1, \cdots, i_k$)分离，则 $\tilde{V}$ 必恰含有 $\boldsymbol{A}$的 k个奇异值.

证明 前一部分结论由引理 2及 V_i 的定义获证. 后一部分结论基于下面 3点观察:

(1) 设 $X \geqslant 0$，令 $\boldsymbol{A}(X) = \boldsymbol{D} + X\boldsymbol{B}$，其中 $\boldsymbol{D} = \mathrm{diag}(a_{11}, a_{22}, \cdots, a_{nn})$，$\boldsymbol{B} = \boldsymbol{A} - \boldsymbol{D}$，于是 $\boldsymbol{A}(0) = \boldsymbol{D}$，$\boldsymbol{A}(1) = \boldsymbol{A}$，由 V_i 的定义可验证 $V_i(\boldsymbol{A}(X)) \subseteq V_i(\boldsymbol{A})$ ($0 \leqslant X \leqslant 1$);

(2) $\boldsymbol{A}(X)$的每一奇异值 $^{e}_i(X)$连续地依赖于 X，当 X由 0连续变化到 1时，$^{e}_i(X)$画出非负半实轴上的连续轨迹，起始于某 $|a_{j_i j_i}|$，终止于 $^{e}_i$;

(3) $\boldsymbol{A}(0)$的奇异值，即 a_i ($i = 1, 2, \cdots, n$)满足 $a_i \in V_i(\boldsymbol{A})$.

由标准的连续性论证[1,5]可证得定理 2.

除文 [3] 外，文 [2] 比较仔细、系统地研究了奇异值的估计问题，其中重要结论是

$$\sigma_{\max}(A) \leqslant \sqrt{\|A\|_1 \|A\|_\infty}, \quad \sigma_{\min}(A) \geqslant \min_i \left\{ |a_{ii}| - \frac{1}{2}(r_i + c_i) \right\} \tag{9}$$

以及借助分隔性质的估计方法．下面的数值例子给出几种方法的比较．

例 1　$A=\begin{bmatrix} 10 & 2 \\ -2 & 2 \end{bmatrix}$，奇异值为 $\sigma_1 = 10.325$，$\sigma_2 = 2.325$．由文 [2] 可得估计：$\sigma_1 \in [10.198, 12]$，$\sigma_2 \in [0, 2.828]$．而由定理 2得 $\sigma_1 \in [9.464, 10.472]$，$\sigma_2 \in [1.528, 2.536]$．可见除 σ_1 的下界外，本文的估计都比较好；文 [2]对 σ_2 下界的估计 $\sigma_2 \geqslant 0$是平凡的，据之无法得到条件数上界的估计，用定理 2则可知 $k(A) \leqslant 10.472/1.528 = 6.853$．

例 2　$A=\begin{bmatrix} 6 & 0 & 0 \\ 1 & 10 & 1 \\ 0 & 1 & 3 \end{bmatrix}$，由文 [2] 估计奇异值：$\sigma_1 \in [10, 11.489]$，$\sigma_2 \in [5.5, 6]$，$\sigma_3 \in [2, 3.464]$；而用定理 2可得：$\sigma_1 \in [9.414, 10.449]$，$\sigma_2 \in [5.551, 6.586]$，$\sigma_3 \in [2.697, 3.282]$．除 σ_1的下界及 σ_2 的上界外，本文结果较好，并可获得较好的条件数上界的估计．

注　关于特征值估计还有著名的 Brualdi定理[1]．人们自然会猜想，可否建立奇异值估计的 Brualdi型定理？即猜想弱不可约矩阵 A 的一切奇异值均属于

$$\bigcup_{\gamma \in C(A)} \left\{ z \in \mathbb{C} \;\Big|\; \prod_{P_i \in \gamma} |z - a_i| \leqslant \prod_{P_i \in \gamma} s_i \right\} \tag{10}$$

其中，$C(A)$是 A的有向图中非平凡圈 γ的集合．下面的例子表明，即使 A是不可约的，上述猜想不成立．设不可约矩阵 $A=\begin{bmatrix} 0 & 0.1 & 0 \\ 0 & 0 & 1 \\ 0.1 & 0 & 0 \end{bmatrix}$，$A$的奇异值 $\sigma_1 = 1$不能落入区域 $\left\{ z \in \mathbb{C} \;|\; |z|^3 \leqslant s_1 s_2 s_3 = 0.1 \right\}$．

参 考 文 献

1　Horn R A, Johnson C R. Matrix analysis. New York: Cambridge University Press, 1985. 380～387

2　Horn R A, Johnson C R. Topics in matrix analysis. New York: Cambridge University Press, 1991. 223～231

3　Qi L Q. Some simple estimates for singular values of a matrix. Linear Algebra Appl, 1984, 56: 105～119

4　陈公宁．矩阵理论与应用．北京：高等教育出版社，1990. 234～235

5　孙继广．矩阵扰动分析．北京：科学出版社，1987. 114～115

Brauer-Type Theorems for Estimating Singular Values of a Matrix

Li Saijian[*]　*Li Luoluo*

Abstract　Let $A=(a_{ij}) \in \mathbb{C}^{n \times n}$, $s_i = \max\left\{ \sum_{j \neq i} |a_{ij}|, \sum_{j \neq i} |a_{ji}| \right\}$ and $a_i = |a_{ii}|$ $(i = 1, 2, \cdots, n)$. Every singular value of A is proved to be in the union $\bigcup_{i \neq j} \left\{ z \in \mathbb{C} \;|\; |z - a_i||z - a_j| \leqslant s_i s_j \right\}$. It is shown that a Brauer-type singular value estimate formula need not hold.

Keywords　singular value, inclusion region, condition number

[*] Department of Scientific Computing and Computer Application, Zhongshan University, Guangzhou 510275, China