关于条件数的一个定理

颜世建

(南京师范大学数学与计算机科学学院,210097,江苏,南京)

[摘要] 对正方线性方程组 $Ax=b$ 给出了一种新的条件数,并讨论了它的几何意义.

[关键词] 行列式,正交程度,条件数

[中图分类号]O241, [文献标识码]A, [文章编号]1001-4616(2004)04-0025-03

0 引言

正方线性方程组 $Ax = b$ 的灵敏性的一个精确度量可通过含参数 ε 的方程组(见文[1])

$$(A+\varepsilon F)x(\varepsilon) = b+\varepsilon f \tag{1}$$

来得到,其中 A、$F \in \mathbf{R}^{n\times n}$,$b$、$f \in \mathbf{R}^n$.如果 A 非奇异,则 $Ax = b$ 有唯一解 x,记为 $x = x(0)$.由隐函数存在定理.A 非奇异时,(1) 在零的一个邻域有解 $x(\varepsilon)$ 存在,且 $x(\varepsilon)$ 在该邻域是可微的,并有

$$\dot{x}(0) = A^{-1}[f - Fx]. \tag{2}$$

于是 $x(\varepsilon)$ 有 Taylor 展开

$$x(\varepsilon) = x + \varepsilon\dot{x}(0) + o(\varepsilon^2) = x + \varepsilon A^{-1}[f - Fx] + o(\varepsilon^2). \tag{3}$$

利用 $\|b\| \leqslant \|A\|\|x\|$ 可得

$$\frac{\|x(\varepsilon)-x\|}{\|x\|} \leqslant \|A\|\|A^{-1}\|\left[\frac{\|\varepsilon f\|}{\|b\|}+\frac{\|\varepsilon F\|}{\|A\|}\right]+o(\varepsilon^2). \tag{4}$$

通常称 $\|A\|\|A^{-1}\|$ 为解 $Ax = b$ 时的条件数,它量化了 $Ax = b$ 的灵敏性.可是 $\|A\|\|A^{-1}\|$ 中含有 A^{-1},这使人略感不便.本文将给出度量 $Ax = b$ 灵敏性的另一种量.它是由 A 的行向量正交程度来刻划的.

1 新的条件数

设 $\dot{x}(0) = (\dot{x}_1(0), \dot{x}_2(0), \cdots, \dot{x}_n(0))^T$, $f = (f_1, f_2, \cdots, f_n)^T$.记

$$F = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{pmatrix}, \quad A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix}$$

用 $A(j \leftarrow y)$ 表示 A 的第 j 列用 y 代替后得到的矩阵,用 $A\begin{pmatrix} i_1, i_2, \cdots, i_s \\ j_1, j_2, \cdots, j_t \end{pmatrix}$ 表示由处于 A 的第 $i_1, i_2, \cdots, i_s$ 行第 $j_1, j_2, \cdots, j_t$ 列的元素组成的子矩阵,由 $A\dot{x}(0) = f - Fx$ 得

$$\dot{x}_j(0) = \frac{\det A(j \leftarrow f - Fx)}{\det A} = \frac{1}{\det A}\sum_{i=1}^{n}(f_i - F_i x)(-1)^{(i+j)}\det A\begin{pmatrix} 1,\cdots,i-1,i+1,\cdots,n \\ 1,\cdots,j-1,j+1,\cdots,n \end{pmatrix}. \tag{5}$$

由矩阵积的行列式展开定理知(见文[2])

$$\sum_{j=1}^{n}\left[\det A\begin{pmatrix} 1,\cdots,i-1,i+1,\cdots,n \\ 1,\cdots,j-1,j+1,\cdots,n \end{pmatrix}\right]^2$$
$$= \sum_{j=1}^{n}\det A\begin{pmatrix} 1,\cdots,i-1,i+1,\cdots,n \\ 1,\cdots,j-1,j+1,\cdots,n \end{pmatrix}\det\begin{pmatrix} 1,\cdots,i-1,i+1,\cdots,n \\ 1,\cdots,j-1,j+1,\cdots,n \end{pmatrix}^T$$

收稿日期:2003-04-16.

作者简介:颜世建,1946 - ,南京师范大学数学与计算机科学学院教授,主要从事计算数学的教学与研究,E - mail:yanshijian@163.com

$$= \det\left[A\begin{pmatrix}1,\cdots,i-1,i+1,\cdots,n\\1,\cdots,j-1,j,j+1,\cdots,n\end{pmatrix}A\begin{pmatrix}1,\cdots,i-1,i+1,\cdots,n\\1,\cdots,j-1,j,j+1,\cdots,n\end{pmatrix}^T\right]$$

再由 Hadamard 不等式知

$$\sum_{j=1}^{n}\left[\det A\begin{pmatrix}1,\cdots,i-1,i+1,\cdots,n\\1,\cdots,j-1,j+1,\cdots,n\end{pmatrix}\right]^2 = \det\left[\begin{pmatrix}A_1\\\vdots\\A_{i-1}\\A_{i+1}\\\vdots\\A_n\end{pmatrix}\begin{pmatrix}A_1\\\vdots\\A_{i-1}\\A_{i+1}\\\vdots\\A_n\end{pmatrix}^T\right] \leqslant \prod_{\substack{j=1\\j\neq i}}^{n}\|A_j\|_2^2. \tag{6}$$

另外由 Cauchy-Schwarz 不等式及(6) 知

$$\left|\sum_{j=1}^{n}\det A\begin{pmatrix}1,\cdots,i-1,i+1,\cdots,n\\1,\cdots,j-1,j+1,\cdots,n\end{pmatrix}\det\begin{pmatrix}1,\cdots,k-1,k+1,\cdots,n\\1,\cdots,j-1,j+1,\cdots,n\end{pmatrix}\right|$$

$$\leqslant\left[\sum_{j=1}^{n}\left(\det A\begin{pmatrix}1,\cdots,i-1,i+1,\cdots,n\\1,\cdots,j-1,j+1,\cdots,n\end{pmatrix}\right)^2\right]^{\frac12}\left[\sum_{j=1}^{n}\left(\det A\begin{pmatrix}1,\cdots,k-1,k+1,\cdots,n\\1,\cdots,j-1,j+1,\cdots,n\end{pmatrix}\right)^2\right]^{\frac12}$$

$$\leqslant\left(\prod_{\substack{j=1\\j\neq n}}^{n}\|A_j\|_2\right)\left(\prod_{\substack{j=1\\j\neq k}}^{n}\|A_j\|_2\right) \tag{7}$$

于是由(5),(6),(7) 得

$$\sum_{j=1}^{n}\dot{x}_j(0)^2 = \frac{1}{(\det A)^2}\sum_{j=1}^{n}\left[\sum_{i=1}^{n}(f_i-F_ix)(-1)^{(i+j)}\det A\begin{pmatrix}1,\cdots,i-1,i+1,\cdots,n\\1,\cdots,j-1,j+1,\cdots,n\end{pmatrix}\right]^2$$

$$= \frac{1}{(\det A)^2}\left\{\sum_{i=1}^{n}(f_i-F_ix)^2\sum_{j=1}^{n}\left[\det A\begin{pmatrix}1,\cdots,i-1,i+1,\cdots,n\\1,\cdots,j-1,j+1,\cdots,n\end{pmatrix}\right]^2\right.$$

$$+\sum_{i=1}^{n}\sum_{k=1}^{n}(f_i-F_ix)(f_k-F_kx)(-1)^{(i+k)}\cdot$$

$$\left.\sum_{j=1}^{n}\det A\begin{pmatrix}1,\cdots,i-1,i+1,\cdots,n\\1,\cdots,j-1,j+1,\cdots,n\end{pmatrix}\det A\begin{pmatrix}1,\cdots,k-1,k+1,\cdots,n\\1,\cdots,j-1,j+1,\cdots,n\end{pmatrix}\right\}$$

$$\leqslant \frac{1}{(\det A)^2}\left[\sum_{i=1}^{n}|f_i-F_ix|^2\left(\prod_{j\neq i}\|A_j\|_2^2\right)\right.$$

$$\left.+\sum_{i=1}^{n}\sum_{k=1}^{n}|f_i-F_ix||f_k-F_kx|\left(\prod_{j\neq i}\|A_j\|_2\right)\left(\prod_{j\neq k}\|A_j\|_2\right)\right]$$

$$= \frac{1}{(\det A)^2}\left[\sum_{i=1}^{n}|f_i-F_ix|\left(\prod_{j\neq i}\|A_j\|_2\right)\right]^2$$

$$= \frac{\prod_{j=1}^{n}\|A_j\|_2^2}{(\det A)^2}\left[\sum_{i=1}^{n}\frac{|f_i|}{\|A_i\|_2}+\sum_{i=1}^{n}\frac{\|F_i\|_2\|x\|_2}{\|A_i\|_2}\right]^2,$$

由(3) 得

$$\frac{\|x(\varepsilon)-x\|_2}{\|x\|_2} \leqslant \frac{\prod_{j=1}^{n}\|A_j\|_2}{|\det A|}\left[\sum_{i=1}^{n}\frac{|\varepsilon f_i|}{\|A_i\|_2\|x\|_2}+\sum_{i=1}^{n}\frac{\|\varepsilon F_i\|_2}{\|A_i\|_2}\right]+o(\varepsilon^2)$$

$$\leqslant \frac{\prod_{j=1}^{n}\|A_j\|_2}{|\det A|}\left[\sum_{i=1}^{n}\frac{|\varepsilon f_i|}{|b_i|}+\sum_{i=1}^{n}\frac{\|\varepsilon F_i\|_2}{\|A_i\|_2}\right]+o(\varepsilon^2).$$

由上面推导可得

定理 1　设 A 非奇异,$b_i\neq 0$,$i=1,\cdots,n$,当 $|\varepsilon|$ 足够小时,$(A+\varepsilon F)x(\varepsilon)=b+\varepsilon f$ 有唯一解 $x(\varepsilon)$. 记 $x=x(0)$,则

$$\frac{\|x(\varepsilon)-x\|_2}{\|x\|_2} \leqslant \frac{\prod_{j=1}^{n}\|A_j\|_2}{|\det A|}\left[\sum_{i=1}^{n}\frac{|\varepsilon f_i|}{|b_i|}+\sum_{i=1}^{n}\frac{\|\varepsilon F_i\|_2}{\|A_i\|_2}\right]+o(\varepsilon^2). \tag{8}$$

下面说明一下$\frac{\prod_{j=1}^{n}\|A_j\|_2}{|\det A|}$的几何意义.

设 $q_1,q_2,\cdots q_k$ 为 k 个 n 维向量,$Q_k=(q_1,q_2,\cdots,q_k)\in R^{n\times k}$.若 $q_1,q_2,\cdots,q_k$ 两两正交,则以 $q_1,q_2,\cdots,q_k$ 为边的超长方体体积$\tilde{V}_k=\|q_1\|_2\|q_2\|_2\cdots\|q_k\|_2=\sqrt{\det Q_k^TQ_k}$.如果 $q_1,q_2,\cdots,q_k$ 不是两两正交,则以 $q_1,q_2,\cdots,q_k$ 为边的超平行多面体体积 V_k 越接近$\|q_1\|_2\|q_2\|_2\cdots\|q_k\|_2$,我们就可以认为 $q_1,q_2,\cdots,q_k$ 的正交程度越高,因而我们可用$\frac{V_k}{\|q_1\|_2\|q_2\|_2\cdots\|q_k\|_2}$作为衡量 $q_1,q_2,\cdots,q_k$ 正交程度的一个标准.

当 $k=2$ 时,$V_2^2=\|q_1\|_2^2\|q_2\|_2^2\sin^2\langle q_1,q_2\rangle=\|q_1\|_2^2\|q_2\|_2^2\left[1-\frac{(q_1^Tq_2)^2}{\|q_1\|_2^2\|q_2\|_2^2}\right]=(q_1^Tq_1)(q_2^Tq_2)-(q_1^Tq_2)^2=\det Q_2^TQ_2$,其中 $Q_2=(q_1q_2)$.

假设 $k=i-1$ 时,$V_{i-1}^2=\det Q_{i-1}^TQ_{i-1}$,则 $k=i$ 时

$$\begin{aligned}
V_i^2 &= V_{i-1}^2\left\|\left(I-P_{R(Q_{i-1}),N(Q_{i-1}^T)}\right)q_i\right\|_2^2=\det Q_{i-1}^TQ_{i-1}\left\|\left(I-Q_{i-1}(Q_{i-1}^TQ_{i-1})^{-1}Q_{i-1}^T\right)q_i\right\|_2^2\\
&=(\det Q_{i-1}^TQ_{i-1})\left(q_i^T\left(I-Q_{i-1}(Q_{i-1}^TQ_{i-1})^{-1}Q_{i-1}^T\right)q_i\right)\\
&=\det\begin{pmatrix} Q_{i-1}^TQ_{i-1} & Q_{i-1}^Tq_i\\ 0 & q_i^T\left(I-Q_{i-1}(Q_{i-1}^TQ_{i-1})^{-1}Q_{i-1}^T\right)q_i\end{pmatrix}\\
&=\det\begin{pmatrix} Q_{i-1}^TQ_{i-1} & Q_{i-1}^Tq_i\\ q_i^TQ_{i-1} & q_i^Tq_i\end{pmatrix}=\det Q_i^TQ_i.
\end{aligned}$$

因而 $V_k=\sqrt{\det Q_k^TQ_k}$.由 Hadamard 不等式知$\frac{V_k}{\|q_1\|_2\|q_2\|_2\cdots\|q_k\|_2}\leqslant 1$,故我们可以认为$\frac{\sqrt{\det Q_k^TQ_k}}{\|q_1\|_2\cdots\|q_k\|_2}$越接近 1,则 $q_1,q_2,\cdots,q_k$ 的正交程度越高.

综上所述,(8) 中的$\frac{\|A_1\|_2\|A_2\|_2\cdots\|A_n\|_2}{|\det A|}$表示 A 的行向量$A_1,A_2,\cdots,A_n$ 的正交程度的倒数.可以认为 $A_1,A_2,\cdots,A_n$ 的正交程度越高.即$\frac{\|A_1\|_2\|A_2\|_2\cdots\|A_n\|_2}{|\det A|}$越接近 1.则 $Ax=b$ 性态越好,这和人们的计算经验是一致的.

[参考文献]

[1] Golub G H, Van Loan C F. 矩阵计算[M]. 北京:科学出版社,2001.

[2] 张远达,熊全淹. 线性代数[M]. 北京:人民教育出版社,1962.

A Theorem for Condition Number

Yan Shijian

(School of Mathematics and Computer Science, Nanjing Normal University, 210097, Nanjing, China)

Abstract: A new condition number for square system of linear equtions is presented, and the geometry significance for the new condition number is discussed.

Key words: determinant, orthogonal stage, condition number

[责任编辑:陆炳新]