Linear Systems With Coefficient Matrices Having Fields of Values Not Containing The Origin *

Ren-Cang Li
Department of Mathematics
University of California at Berkeley
Berkeley, California 94720

March 9, 1994

Computer Science Division Technical Report UCB//CSD-94-853, University of California, Berkeley, CA 94720, December, 1994.

This is a continuation of [3] addressing a problem posed by Prof. Kahan. The problem is the following:

> *Given an $n \times n$ (complex) matrix A whose field of values $\mathcal{F}(A)$ does not contain the origin, is it necessary to pivot when solving the linear system $Ax = b$?*

It is well known $\mathcal{F}(A)$ is a compact convex set on the complex plane. Let's draw two projecting lines ℓ_1 and ℓ_2 starting at the origin and "tangent"[1] to the boundary of $\mathcal{F}(A)$ such that $\mathcal{F}(A)$ falls into the smaller section enclosed by ℓ_1 and ℓ_2 as shown in Figure 1. Let α be the angle of the section. Clearly $0 \le \alpha < \pi$. Set $\theta = \pi - \alpha$. It is proved in [3] that if θ is reasonably

*This material is based in part upon work supported by Argonne National Laboratory under grant No. 20552402 and the University of Tennessee through the Advanced Research Projects Agency under contract No. DAAL03-91-C-0047, by the National Science Foundation under grant No. ASC-9005933, and by the National Science Infrastructure grants No. CDA-8722788 and CDA-9401156.

[1] Here "tangent" may not be the right word since the boundary of $\mathcal{F}(A)$ could not be smooth. So to be more rigous, we could say that ℓ_1 and ℓ_2 are two support lines passing through the origin.

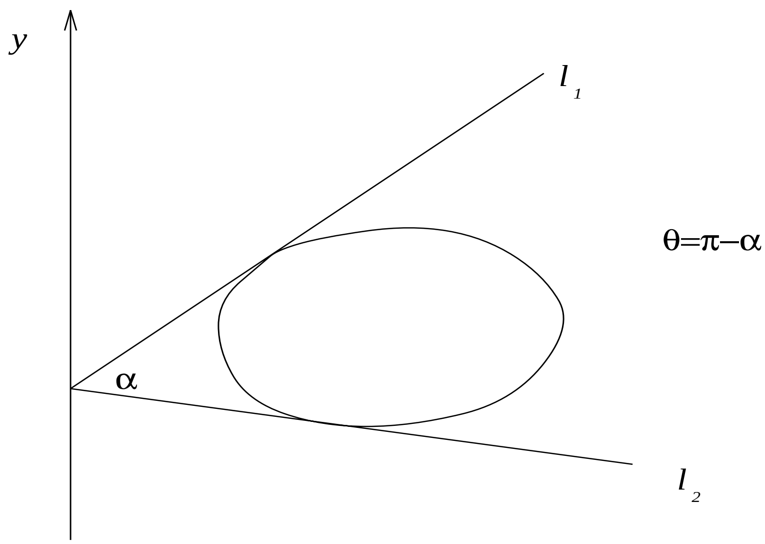

Figure 1: A typical picture.

large, no smaller than 0.1π (say), there is no danger of instability in solving $Ax = b$ without pivoting. However, if θ is rather small, $\theta = \epsilon$ (say), there is a potentiality that elements might grow by a factor $O\left(\frac{1}{\epsilon^3}\right)$. Recently, Ming Gu [2] improves this factor to $O\left(\frac{1}{\epsilon^2}\right)$, which is optimal as far as only order is concerned. This note adopts the idea developed in [1] where the case A being real is studied. We will give a better bound which is *asymptotically attainable*.

It is easy to verify that scalar multiplications do not affect element growth in Gaussian elimination processes. Therefore, without loss of any generality, by rotating the matrix A by an angle β as $e^{i\beta}A$ we can assume that[2] $\mathcal{F}(A)$ lies in the right half plane and the angles between the y-axis and ℓ_1 and between the negative direction of the y-axis and ℓ_2 are equal to $\frac{\theta}{2}$. Set $A = H + iS$, where

$$H = \frac{A + A^*}{2} = H^*,\ S = \frac{A - A^*}{2i} = S^*$$

are both Hermitian, and moreover H is positive definite.

[2] When A is real, $\mathcal{F}(A)$ is symmetric with respect to the x-axis, so either A itself or $-A$ has the desired property.

Doing Gaussian elimination on A, we get a decomposition

$$A = H + iS = LDM^*, \tag{1}$$

where L and M are unit lower triangular matrices, D diagonal. Generally, they are all complex. The existence of the decomposition (1) is guaranteed by the assumption we made on the $\mathcal{F}(A)$ (ref. [3]).

Proposition 1 *Write* $D = \mathrm{diag}\,(d_1, d_2, \cdots, d_n)$*. Then*

$$\Re d_j > 0, \quad j = 1, 2, \cdots, n,$$

where $\Re(\cdot)$ *denotes the real part of a complex number.*

Since $H = H^*$ is positive definite, it has a unique Cholesky decomposition $H = GG^*$, where G is lower triangular. Now (1) gives

$$L^{-1}(H + iS)L^{-*} = DM^*L^{-*} \Rightarrow L^{-1}GG^*L^{-*} + iL^{-1}SL^{-*} = DM^*L^{-*},$$

which yields

$$(G^*L^{-*})^*\, G^*L^{-*} + iL^{-1}SL^{-*} = DM^*L^{-*}. \tag{2}$$

Let e_j be the jth column of the $n \times n$ identity matrix. Comparing the jth diagonal entries of the two sides of (2) leads to

$$\Re d_j = \|G^*L^{-*}e_j\|_2^2 \tag{3}$$

since M^*L^{-*} is unit upper triangular. Therefore[3]

$$\begin{aligned} \|G^*L^{-*}D^{-1/2}\|_F^2 &= \sum_{j=1}^{n} \|G^*L^{-*}D^{-1/2}e_j\|_2^2 \\ &= \sum_{j=1}^{n} \frac{\|G^*L^{-*}e_j\|_2^2}{|d_j|} \\ &\leq n. \end{aligned} \tag{4}$$

On the other hand,

$$M^{-1}(H + iS)M^{-*} = M^{-1}LD \Rightarrow M^{-1}GG^*M^{-*} + iM^{-1}SM^{-*} = M^{-1}LD,$$

[3] $D^{1/2}$ is not single-valued. But for our purpose it is good enough to pick any one of them and stick to it. $D^{-1/2} \stackrel{\mathrm{def}}{=} (D^{1/2})^{-1}$.

which yields $(G^*M^{-*})^* G^*M^{-*} + iM^{-1}SM^{-*} = M^{-1}LD$, so

$$\Re d_j = \|G^*M^{-*}e_j\|_2^2$$

and

$$\|G^*M^{-*}D^{-1/2}\|_F^2 \le n. \tag{5}$$

It follows from (1) that

$$\begin{aligned} LD^{1/2} &= (GG^* + iS)M^{-*}D^{-1/2} \\ &= (G + iSG^{-*})G^*M^{-*}D^{-1/2}, \\ D^{1/2}M &= D^{-1/2}L^{-1}(GG^* + iS) \\ &= D^{-1/2}L^{-1}G(G^* + iG^{-1}S). \end{aligned}$$

Thus

$$\|LD^{1/2}\|_F \le \sqrt{n}\|G + iSG^{-*}\|_2, \tag{6}$$
$$\|D^{1/2}M\|_F \le \sqrt{n}\|G^* + iG^{-1}S\|_2. \tag{7}$$

Notice that

$$\begin{aligned} \|G + iSG^{-*}\|_2^2 &= \|(G + iSG^{-*})(G^* - iG^{-1}S)\|_2 \\ &= \|GG^* - iS + iS + SG^{-*}G^{-1}S\|_2 \\ &= \|H + SH^{-1}S\|_2 \\ &\le \|H\|_2 + \|SH^{-1}S\|_2, \\ \|G^* + iG^{-1}S\|_2^2 &= \|(G - iSG^{-*})(G^* + iG^{-1}S\|_2 \\ &= \|GG^* + iS - iS + SG^{-*}G^{-1}S\|_2 \\ &= \|H + SH^{-1}S\|_2 \\ &\le \|H\|_2 + \|SH^{-1}S\|_2. \end{aligned}$$

Together with (6) and (7), we have[4]

$$\begin{aligned} \|\,|L|\,|D|\,|M^*|\,\|_F &= \|\,|LD^{1/2}|\,|D^{1/2}M|\,\|_F \\ &\le n\|H + SH^{-1}S\|_2 \\ &\le n(\|H\|_2 + \|SH^{-1}S\|_2). \end{aligned} \tag{8}$$

---

[4] By $|X|$, we mean its entrywise absolute value, i.e. $|X| \stackrel{\text{def}}{=} (|x_{ij}|)$.

To relate this bound to the angle α, we observe

$$\begin{aligned}\|SH^{-1}S\|_2 &= \|H^{1/2}(H^{-1/2}SH^{-1/2})(H^{-1/2}SH^{-1/2})H^{1/2}\|_2 \\ &\leq \|H^{1/2}\|_2\|H^{-1/2}SH^{-1/2}\|_2^2\|H^{1/2}\|_2 \\ &= \|H\|_2\|H^{-1/2}SH^{-1/2}\|_2^2.\end{aligned}$$

Lemma 1

$$\|H^{-1/2}SH^{-1/2}\|_2 = \max_{x\neq 0} \frac{x^*Sx}{x^*Hx} = \tan\frac{\alpha}{2}.$$

With those in mind, we get

Theorem 1

$$\|\,|L|\,|D|\,|M^*|\,\|_F \leq n\|H\|_2\left[1+\left(\tan\frac{\alpha}{2}\right)^2\right]. \tag{9}$$

Roughly speaking, the bound in [2] is the one obtained by replacing the number inside $[\cdot]$ of (9) with $1+\tan\frac{\alpha}{2}+\frac{3}{2}\left(\tan\frac{\alpha}{2}\right)^2$.

In what follows, we are going to present an example to shown that this inequality is *at least asymptotically attainable* in the sense that there are examples for which the two sides of (9) are arbitrarily close. Consider (ref. [3])

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 1 & 2r \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1-r & r \\ -r & 1+r \end{pmatrix},$$

where r is positive. It is known the field of values of $\mathcal{F}(A)$ is a disk with center 1 and radius r, i.e.

$$\mathcal{F}(A) = \{z \text{ complex} : |z-1| \leq r\}.$$

So $0 \notin \mathcal{F}(A)$ if $r < 1$ (which will be assumed hereafter). For this A, we have

$$\begin{aligned} A = LDM^* &= \begin{pmatrix} 1 & 0 \\ -\frac{r}{1-r} & 1 \end{pmatrix}\begin{pmatrix} 1-r & 0 \\ 0 & \frac{1}{1-r} \end{pmatrix}\begin{pmatrix} 1 & \frac{r}{1-r} \\ 0 & 1 \end{pmatrix}, \\ |L|\,|D|\,|M^*| &= \begin{pmatrix} 1 & 0 \\ \frac{r}{1-r} & 1 \end{pmatrix}\begin{pmatrix} 1-r & 0 \\ 0 & \frac{1}{1-r} \end{pmatrix}\begin{pmatrix} 1 & \frac{r}{1-r} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-r & r \\ r & \frac{1+r^2}{1-r} \end{pmatrix}, \\ H &= \frac{A+A^*}{2} = \begin{pmatrix} 1-r & 0 \\ 0 & 1+r \end{pmatrix}, \end{aligned}$$

$$
\begin{aligned}
S &= \frac{A - A^*}{2i} = \begin{pmatrix} 0 & -ri \\ ri & 0 \end{pmatrix}, \\
SH^{-1}S &= \begin{pmatrix} \frac{r^2}{1-r} & 0 \\ 0 & \frac{r^2}{1+r} \end{pmatrix}, \\
\tan\frac{\alpha}{2} &= \frac{r}{\sqrt{1-r^2}}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
\|\,|L|\,|D|\,|M^*|\,\|_F^2 &= (1-r)^2 + 2r^2 + \left(\frac{1+r^2}{1-r}\right)^2 \\
&= \frac{(1-r)^4 + 2r^2(1-r)^2 + (1+r^2)^2}{(1-r)^2}, \\
2\|H\|_2 \left[1 + \left(\tan\frac{\alpha}{2}\right)^2\right] &= 2(1+r)\left[1 + \frac{r^2}{1-r^2}\right] = \frac{2}{1-r}, \\
2(\|H\|_2 + \|SH^{-1}S\|_2) &= \frac{2}{1-r}.
\end{aligned}
$$

Define a function $f(r)$ as follows:

$$
\begin{aligned}
f(r) &\stackrel{\text{def}}{=} \frac{\|\,|L|\,|D|\,|M^*|\,\|_F}{2\|H\|_2\left[1 + (\tan\frac{\alpha}{2})^2\right]} \\
&= \frac{1}{2}\sqrt{(1-r)^4 + 2r^2(1-r)^2 + (1+r^2)^2} \\
&= 1 - (1-r) + \frac{3}{4}(1-r)^2 - \frac{1}{4}(1-r)^3 - \frac{1}{32}(1-r)^4 + O((1-r)^5).
\end{aligned}
$$

It is easy to see $f(0) = 1/\sqrt{2} = 0.70710678118655$, $\lim_{r\to 1^-} f(r) = 1 = f(1)$ which shows that the inequalities (8) and (9) are *asymptotically attainable!* And $\min_{0\le r\le 1} f(r) \approx 0.614966762630915$ at

$$
r = \frac{1}{2} - \frac{\sqrt[3]{2/3}}{\sqrt[3]{-9+\sqrt{177}}} + \frac{\sqrt[3]{-9+\sqrt{177}}}{\sqrt[3]{12^2}} \approx 0.273301174242.
$$

To see how fast $f(r)$ approaches 1 pictorially, we refer the reader to Figure 2, where the picture on the left is the graph of $f(r)$ and the one on the right is that of $1 - f(r)$.

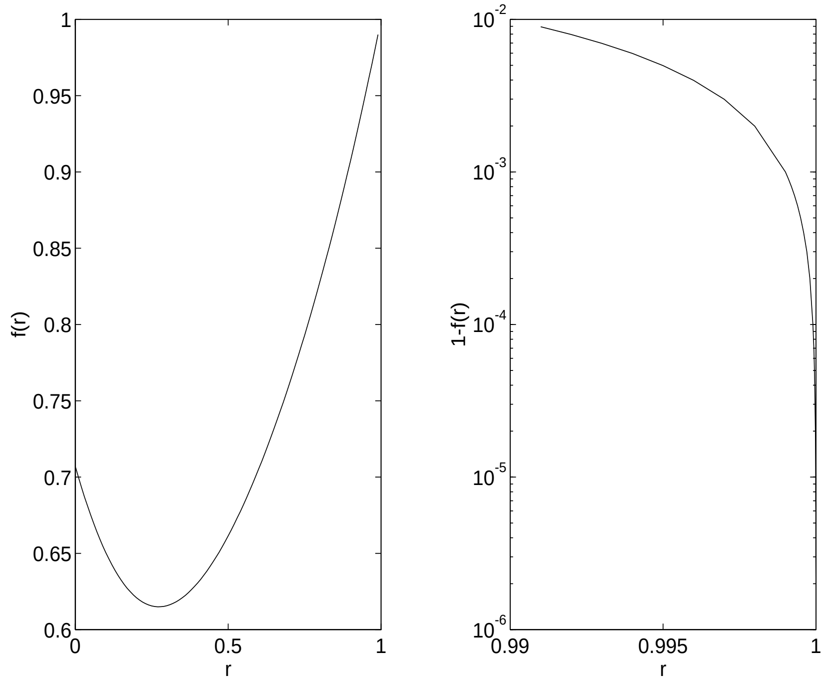

Figure 2: The functions $f(r)$ and $1 - f(r)$.

References

[1] G. H. Golub and Ch. van Loan, Unsymmetric positive definite linear systems, *Linear Algebra and its Applications*, **28**(1979), 85–97.

[2] M. Gu, Relating element growth of Gaussian elimination to the field of values, 1994.

[3] R.-C. Li, Relations between the field of values of a matrix and those of its Schur complements, preprint, 1992.