Lower and Upper Bounds in the Perturbation of General Linear Algebraic Equations

JIU DING
Department of Mathematics, The University of Southern Mississippi
Hattiesburg, MS 39406-5045, U.S.A.

(Received December 1999; accepted January 2000)

Communicated by R. P. Srivastav

Abstract—A perturbation result concerning the upper and lower bounds on relative errors of solutions to linear algebraic equations is given under any norm for consistent systems and under the Euclidean norm for general systems. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords—Least squares solution, Generalized inverse.

1. INTRODUCTION

Consider the nonhomogeneous system of linear algebraic equations

\[ Ax = b, \]

where \( A \in \mathbb{R}^{m \times n} \) is an \( m \times n \) real matrix and \( b \in \mathbb{R}^m \) is an \( m \)-dimensional real vector. Let (1) be perturbed to

\[ (A + E)y = b + e. \]

Perturbation analysis for solving linear equations (1) is of practical importance in numerical algebra and scientific computing [1-3]. A classical perturbation result (see, e.g., [2]) states that if \( A \) is nonsingular, then for the exact solution \( x = A^{-1}b \) and an approximate solution \( \hat{x} \) of (1),

\[ \frac{1}{\kappa} \cdot \frac{\| r \|}{\| b \|} \leq \frac{\| x - \hat{x} \|}{\| x \|} \leq \kappa \cdot \frac{\| r \|}{\| b \|}, \]

where \( \kappa = \| A \| \| A^{-1} \| \) is the condition number of \( A \). It is well known that the reciprocal of the condition number measures how near the given nonsingular problem (1) is to singularity [4], and this is also reflected in the lower bound of the relative error in (3). In this short note, we want to extend the above result, which only concerns a special perturbation of \( E = 0 \), to general matrices and general perturbations. We will give lower bounds as well as upper bounds of relative errors.

Research was supported in part by an Aubrey K. Lucas Endowment award from the University of Southern Mississippi.

0893-9659/00/$ - see front matter ©2000 Elsevier Science Ltd. All rights reserved. Typeset by A\LaTeX

of perturbed solutions. Such results seem new, simple, and general, and are reduced to known ones [1,2] in special cases. In the next section, we consider the case when (1) and (2) are both consistent, and the more general case of least squares problems will be dealt with in Section 3.

2. CONSISTENT SYSTEMS

In this section, we assume that the norm \( \| \cdot \| \) for \( \mathbb{R}^n \) and \( \mathbb{R}^m \) is arbitrary, and the matrix norm on \( \mathbb{R}^{m \times n} \) is the induced one from the norms on \( \mathbb{R}^n \) and \( \mathbb{R}^m \). Since the matrix \( A \in \mathbb{R}^{m \times n} \) is arbitrary, we need the concept of generalized inverses of matrices for the perturbation analysis.

Let \( N(A) \) and \( R(A) \) be the null space and the range of \( A \), respectively. Let \( P \in \mathbb{R}^{m \times n} \) and \( Q \in \mathbb{R}^{m \times m} \) be two projections (i.e., \( P^2 = P \) and \( Q^2 = Q \)) such that

\[
R(P) = N(A), \quad N(P) = N(A)^c, \quad R(Q) = R(A), \quad N(Q) = R(A)^c,
\]

where \( R^n = N(A) \oplus N(A)^c \) and \( R^m = R(A) \oplus R(A)^c \) are two direct sum decompositions. Since \( A \) maps \( N(A)^c \) to \( R(A) \) bijectively, the generalized inverse \( A^\dagger \in \mathbb{R}^{m \times n} \) of \( A \) associated with \( P, Q \) is defined by

\[
A^\dagger y = x
\]

for \( y \in R(A) \), and \( A^\dagger y = 0 \) for \( y \in R(A)^c \). \( A^\dagger \) is the unique \( n \times m \) matrix satisfying

\[
AA^\dagger A = A, \quad A^\dagger AA^\dagger = A^\dagger, \quad A^\dagger A = I - P, \quad AA^\dagger = Q.
\]

When the usual Euclidean norm \( \| \cdot \|_2 \) is used for \( \mathbb{R}^n \) and \( \mathbb{R}^m \), we require further that \( P \) and \( Q \) are self-adjoint, so that \( A^\dagger A \) and \( AA^\dagger \) are orthogonal projections. In this case, \( A^\dagger \) is called the Moore-Penrose generalized inverse of \( A \). See [5] or [6] for more details on \( A^\dagger \). The Moore-Penrose generalized inverse will be used in Section 3 when we study least squares problems.

Let \( \kappa = \| A \| \| A^\dagger \| \) be the condition number of \( A \) with respect to the induced matrix norm, and denote by \( d(x, N(A)) \) the distance of \( x \) to \( N(A) \) under the vector norm.

**Theorem 2.1.** Suppose \( b \in R(A) \) and \( b + e \in R(A + E) \). If \( \| A^\dagger \| \| E \| < 1 \), then for any solution \( y \) to (2), there is a solution \( x \) to (1) such that

\[
\frac{1}{\kappa + 1} \frac{\| e - Ex \|}{\| b \|} \leq \frac{\| y - x \|}{d(x, N(A))} \leq \frac{\kappa}{1 - \| A^\dagger E \|} \frac{\| e - Ex \|}{\| b \|}.
\]

**Proof.** Let \( x = A^\dagger b + (I - A^\dagger A)y \) be the projection of \( y \) onto the solution set of (1) along \( N(A) \). Subtracting \( Ax = b \) from \( (A + E)y = b + e \) gives that

\[
A(y - x) + Ey = e,
\]

from which we obtain

\[
(A + E)(y - x) = e - Ex.
\]

Multiplying \( A^\dagger \) from left to (5) and using the fact that \( y - x = A^\dagger A(y - x) \) and \( (I + A^\dagger E)^{-1} \) exists, we have

\[
y - x = (I + A^\dagger E)^{-1} A^\dagger (e - Ex).
\]

Let \( z \in N(A) \) be arbitrary. Then (6) gives that

\[
\frac{\| y - x \|}{\| x - z \|} \leq \frac{1}{1 - \| A^\dagger E \|} \frac{\| e - Ex \|}{\| x - z \|} = \frac{\| A \| \| A^\dagger \| \| e - Ex \|}{1 - \| A^\dagger E \| \| x - z \|} \leq \frac{\kappa}{1 - \| A^\dagger E \|} \frac{\| e - Ex \|}{\| A(x - z) \|} = \frac{\kappa}{1 - \| A^\dagger E \|} \frac{\| e - Ex \|}{\| b \|},
\]

from which it follows that

\[
\frac{\| y - x \|}{d(x, N(A))} \leq \frac{\kappa}{1 - \| A^\dagger E \|} \frac{\| e - Ex \|}{\| b \|}.
\]
Now (5) implies that
\[
\frac{\|e - Ex\|}{\|b\|} \leq \frac{\|A + E\|}{\|b\|} \|y - x\| \leq \frac{\|A\|}{\|A\|} (\|A\| + \|E\|) \frac{\|y - x\|}{\|b\|} \\
\leq \frac{(\kappa + 1)\|y - x\|}{\|A^t b\|} \leq \frac{(\kappa + 1)\|y - x\|}{\|A^t b\|}.
\]
Hence, the left inequality of (4) is also true.

**Remark 2.1.** When only \(\|\cdot\|_2\) is used for the vector norm, \(\|y - x\|_2\) is the Euclidean distance of \(y\) to the solution set of (1), and in this case (4) gives lower and upper bounds of this distance with respect to the distance of the solution \(x\) to the null space of \(A\).

**Corollary 2.1.** If \(A\) is nonsingular and \(\|A^{-1}\|\|E\| < 1\), then
\[
\frac{1}{\kappa + 1} \frac{\|e - Ex\|}{\|b\|} \leq \frac{\|y - x\|}{\|x\|} \leq \frac{\kappa}{1 - \|A^{-1}E\|} \frac{\|e - Ex\|}{\|b\|},
\]
where \(x = A^{-1}b\) and \(y = (A + E)^{-1}(b + e)\). In particular, if \(E = 0\), then
\[
\frac{1}{\kappa} \frac{\|e\|}{\|b\|} \leq \frac{\|y - x\|}{\|x\|} \leq \kappa \cdot \frac{\|e\|}{\|b\|}.
\]

**Remark 2.2.** Equation (8) is just (3); see also (2.3.3') and (2.3.11) in [3].

### 3. Least Squares Problems

Now we get rid of the consistency assumption in Theorem 2.1. Let \(A \in \mathbb{R}^{m \times n}\) and \(b \in \mathbb{R}^m\), and consider the least squares problem
\[
\|Ax - b\| = \min_{z \in \mathbb{R}^n} \|Az - b\|,
\]
where the norm \(\|\cdot\| \equiv \|\cdot\|_2\). Suppose (9) is perturbed to
\[
\|(A + E)y - (b + e)\| = \min_{z \in \mathbb{R}^n} \|(A + E)z - (b + e)\|.
\]

An extensive review on perturbation results of rank deficient least squares problems was given in [1] and a recent result was obtained in [7]. Here, we follow the approach in the previous section, which originated from the idea in [8]. We need the following lemma [6, Theorem 8.5].

**Lemma 3.1.**
\[
A^t - (A + E)^t \quad \text{is a symmetric \(m \times m\) matrix.}
\]

**Theorem 3.1.** If \(\|A\|\|E\| < 1\), then for any solution \(y\) to (10), there is a solution \(x\) to (9) such that
\[
\frac{1}{\kappa + 1} \frac{\|e - Ex + r' - r\|}{\|AA^t b\|} \leq \frac{\|y - x\|}{d(x, N(A))} \leq \frac{\kappa}{1 - \|A^{-1}E\|} \frac{\|E\|}{\|A^t b\|} + \frac{2\|e - Ex\|}{\|AA^t b\|},
\]
with \(r = Ax - b\) and \(r' = (A + E)y - (b + e)\).

**Proof.** Let \(x = A^t b + (I - A^t A)y\) be the orthogonal projection of \(y\) onto the solution set of (9). Then \(r = Ax - b\) and \(r' = (A + E)y - (b + e)\) give that
\[
(A + E)(y - x) = e - Ex + r' - r.
\]
Multiplying $A^\dagger$ to (12) from left and using the fact that $y - x = A^\dagger A(y - x)$, $A^\dagger r = 0$, and $(A + E)^\dagger r' = 0$, we have

$$y - x = (I + A^\dagger E)^{-1} A^\dagger (e - Ex) + [A^\dagger - (A + E)^\dagger] r'.$$

By Lemma 3.1,

$$[A^\dagger - (A + E)^\dagger] r' = -A^\dagger (A^\dagger)^T E^T r',$$

from which and the fact that $\|(EA^\dagger)^T\| = \|EA^\dagger\|$, it follows that

$$\|\{A^\dagger - (A + E)^\dagger\} r'\| \leq \|A^\dagger\| \|EA^\dagger\| \|r'\|.

(14)

Moreover, since $y$ solves (10),

$$\|r'\| \leq \|(A + E)x - (b + e)\| \leq \|r\| + \|e - Ex\|.

(15)

Therefore, by (13)-(15), we have, for any $z \in N(A)$,

$$\frac{\|y - x\|}{\|x - z\|} \leq \frac{\|A^\dagger\| \|EA^\dagger\| \|r'\| + \|e - Ex\|}{1 - \|A^\dagger E\| \|x - z\|} \leq \frac{\|A\| \|A^\dagger\| \|EA^\dagger\| \|r\| + \|e - Ex\|}{1 - \|A^\dagger E\| \|Ax\|} \leq \frac{\kappa \|EA^\dagger\| \|r\| + \|e - Ex\|}{1 - \|A^\dagger E\| \|AA^\dagger b\|},$$

which gives the right inequality of (11). The left one in (11) is from

$$\frac{\|e - Ex + r' - r\|}{\|AA^\dagger b\|} \leq \frac{\|A + E\| \|y - z\|}{\|AA^\dagger b\|} \leq \frac{(\kappa + \|A^\dagger\| \|E\|) \|y - x\|}{\|AA^\dagger b\|} \leq \frac{(\kappa + 1)\|y - x\|}{\|x - (I - A^\dagger A) y\|} \leq \frac{(\kappa + 1)\|y - x\|}{d(x, N(A))}.$$

**Corollary 3.1.** If in addition $b \in R(A)$, then

$$\frac{1}{\kappa + 1} \frac{\|e - Ex + r'\|}{\|b\|} \leq \frac{\|y - x\|}{d(x, N(A))} \leq \frac{\kappa}{1 - \|A^\dagger E\|} \frac{2\|e - Ex\|}{\|b\|}.

(16)

If both $b \in R(A)$ and $b + e \in R(A + E)$, then (11) is reduced to (4).

**Corollary 3.2.** If in addition $A$ is of full column rank, then $x = A^\dagger b$ and $y = (A + E)^\dagger (b + e)$, and

$$\frac{1}{\kappa + 1} \frac{\|e - Ex + r' - r\|}{\|AA^\dagger b\|} \leq \frac{\|y - x\|}{\|x\|} \leq \frac{\kappa \|EA^\dagger\| \|r\| + 2\|e - Ex\|}{1 - \|A^\dagger E\| \|AA^\dagger b\|}.$$

(17)

Finally, we mention that Theorem 2.1 is still true for bounded linear operators with closed range between general Banach spaces such that the generalized inverse is well defined, and Theorem 3.1 can be directly generalized for bounded linear operators of Hilbert spaces.

**REFERENCES**