Computing 19, 149—155 (1977)

Computing

A Measure for Ill-Conditioning of Matrices in Interval Arithmetic

J. Zlámal, Ostrava-Kunčice

Received December 5, 1975; revised January 27, 1977

Abstract — Zusammenfassung

A Measure for Ill-Conditioning of Matrices in Interval Arithmetic. The new condition number $V(A)$ is proposed. This condition number measures ill-conditioning in interval arithmetic. The reliability of the condition number $V(A)$ has been proved. It is shown that Bauer's minimum condition number [2] and the condition number $V(A)$ are essentially equivalent although different approaches were used to derive them.

Ein Maß für die Kondition von Matrizen bei Intervallarithmetik. Es wird ein neues Konditionsmaß $V(A)$ vorgeschlagen, das die Kondition von Matrizen in der Intervallarithmetik angibt. Im experimentellen Teil wird die Verläßlichkeit dieses Konditionsmaßes $V(A)$ gezeigt. Es ist nachgewiesen, daß die minimale Konditionszahl nach Bauer [2] und das Konditionsmaß $V(A)$ im wesentlichen äquivalent sind, obwohl zu ihrer Herleitung grundverschiedene Überlegungen benützt werden.

1. Introduction

One often deals with matrices which have elements known only to a certain degree of accuracy. In such cases it is useful to use interval arithmetic.

It is desirable to know the ill-conditioning of a matrix when inverting a matrix or when a system of linear algebraic equations is to be solved. Knowing the ill-conditioning we can estimate the precision of the computed inversion matrix or the precision of the solution of a system of linear algebraic equations.

The goal of this paper is to propose a measure of ill-conditioning in matrix interval arithmetic and to verify the reliability of this measure. It is claimed that the proposed measure is essentially equivalent with Bauer's minimum condition number [2]. A function was found for the estimation of the proposed measure using Bauer's minimum condition number.

2. The Condition Number V(A)

In interval arithmetic a number is characterized by a low bound and a high bound. The notation is as follows:

$$a := [a_1, a_2], \text{ where } a_1 \leqq a \leqq a_2.$$

Four basic arithmetic operations are well known and are defined in [1].

Let A^I be a matrix whose elements are interval numbers. The inversion matrix $(A^I)^{-1}$ is defined as the smallest interval matrix containing the solution set

$$\{A^{-1} \mid A \in A^I, A A^{-1} = I\}.$$

For simplicity, assume all $A \in A^I$ are nonsingular. The notation $A \in A^I$ indicates that each element of the noninterval matrix A is contained in the corresponding interval element of A^I.

An approximation of $(A^I)^{-1}$ can be computed by using a finite method for matrix inversion but performing arithmetic operations in interval arithmetic. It is known that for a general matrix A^I such a method suffers from an inherent loss of accuracy. The interval elements of an approximation of an inversion matrix $(A^I)^{-1}$ will be relatively large.

To verify this I inverted a matrix with order $n=40$ by the Gauss method performed in interval arithmetic. The interval elements of the inversion matrix were on the average ten times larger than the interval elements of the original matrix. We can assume that this ratio depends on the ill-conditioning of the matrix.

Eldon Hansen [3] has proposed the method which minimizes the inherent error of interval arithmetic. This method was used in the experimental part of this paper. The basic idea of this method is as follows:

Let $A_c \in A^I$ be a matrix with elements which are "the centres" of corresponding intervals. Using ordinary arithmetic compute $B = A_c^{-1}$. In interval arithmetic compute $A^I B$. Since $A^I B = I - E^I$, where the interval elements of E^I are "small" it is, in general, possible to express

$$(A^I B)^{-1} = I + E^I + (E^I)^2 + \dots . \tag{1}$$

Hence one can obtain $(A^I B)^{-1}$ by truncating this series. We can rewrite $(A^I)^{-1}$ as $B(A^I B)^{-1}$. The errors inherent in inverting A^I by interval arithmetic are replaced by the errors in inverting $A^I B$. The latter errors are considerably smaller. For an interval matrix A^I, this method computes an interval matrix C^I containing $(A^I)^{-1}$. In practice the length of each interval element of C^I can be made relatively small so that "good" bounds for the elements of A^{-1} are obtained. See [3].

Is there a great difference between C^I and $(A^I)^{-1}$? Moore showed that there exists a bound on the difference between C^I and $(A^I)^{-1}$ which is of second order in the widths of the elements of A^I. See [1] page 57. In the following we will assume that such second order terms and also rounding errors are negligible.

If we know an interval inversion matrix the question arises how to evaluate this matrix with respect to the original inversion matrix. This will be important if the order of the matrix is high.

We can generally assume that the error associated with input data is directly proportional to the absolute value of that data in standard applications. This assumption suggests that we might express interval elements in relation to the "centres" of interval elements. In fact it is the principle of relative error. Since there exist great differences among corresponding elements of a matrix and an inversion matrix I propose the evaluation of the interval inversion matrix by means of percentage. That is, low and high bounds can be computed for each element of A^I as follows:

$$A^I=(a^I_{ij}), \quad a^I_{ij}:=[a^{(1)}_{ij}, a^{(2)}_{ij}],$$

$$a^{(1)}_{ij}=(100-q^{(1)}_{ij})\, a_{ij}/100, \tag{2}$$

$$a^{(2)}_{ij}=(100+q^{(2)}_{ij})\, a_{ij}/100, \tag{3}$$

where a_{ij} is the "centre" of the interval element a^I_{ij}. If $a_{ij}=0$ then we set $a^I_{ij}:=[0, 0]$ and define $q^{(1)}_{ij}=0$, $q^{(2)}_{ij}=0$.

We can express low and high bounds of an interval inversion matrix in similar manner:

$$(A^I)^{-1}=(\alpha^I_{ij}), \quad \alpha^I_{ij}:=[\alpha^{(1)}_{ij}, \alpha^{(2)}_{ij}],$$

$$\alpha^{(1)}_{ij}=(100-p^{(1)}_{ij})\, \alpha_{ij}/100, \tag{4}$$

$$\alpha^{(2)}_{ij}=(100+p^{(2)}_{ij})\, \alpha_{ij}/100, \tag{5}$$

where α_{ij} is the "centre" of the interval element α^I_{ij}. If $\alpha_{ij}=0$ then we set $\alpha^I_{ij}:=[0, 0]$ and define $p^{(1)}_{ij}=0$, $p^{(2)}_{ij}=0$.

The measure for ill-conditioning of an interval matrix — the new condition number — can be defined:

$$V(A)=\frac{p}{q},$$

where

$$p=\frac{\sum_{i,j} \max\,(p^{(1)}_{ij}, p^{(2)}_{ij})}{n^2},$$

$$q=\frac{\sum_{i,j} \max\,(q^{(1)}_{ij}, q^{(2)}_{ij})}{n^2}, \quad i,j=1,2,\ldots,n.$$

3. The Justification of the Condition Number V(A)

To justify the condition number $V(A)$ I have used a battery of test matrices with inversion matrices known explicitly. This battery represents matrices of all important types, e.g. symmetric, non symmetric, well conditioned, extremely ill-conditioned, positively definite, negatively definite, sparse matrices, matrices without zero element and quasidiagonal. The battery is composed of 16 test matrices. The description of the battery can be found in [5].

Each test matrix was considered as an A_c matrix. For every element a_{ij} the interval element a^I_{ij} was computed according to the formula:

$$a^{(1)}_{ij}=(100-q)\, a_{ij}/100,$$

$$a^{(2)}_{ij}=(100+q)\, a_{ij}/100.$$

This interval matrix was inverted by the Hansen method. Earlier we defined $A_c^{-1}=B$. The elements of matrix B were considered as the "centres" of corresponding interval elements of the matrix $(A^I)^{-1}$.

The factors $p_{ij}^{(1)}$, $p_{ij}^{(2)}$, $i,j=1,2,\dots,n$ were computed according to (4) and (5).

The condition number $V(A)$ can be computed now:

$$V(A)=\frac{p}{q},$$

where q is a constant in this case.

Two variants were computed:

A for the order of well-conditioned matrices,
B for the order of ill-conditioned matrices.

$q=0.1$

$A: n=5, 10, 15, 20, 25, 30, 35,$
$B: n=5, 6, 7, 8, 9, 10, 11.$

$q=0.2, 0.4, 1$

$A: n=5, 10, 15,$
$B: n=5, 6, 7.$

To assure convergence according to (1) the norm of E^I matrix must be lower than 1. I have used the following norm for the interval matrix A^I:

$$\| A^I \| = \max_i \sum_j \max (|a_{ij}^{(1)}|, |a_{ij}^{(2)}|).$$

This condition was not fulfilled in many cases especially by ill-conditioned matrices. The following value was then used for q: 0.000001. If this value did not satisfy the condition the interval inversion matrix was not computed. This occurred when computing ill-conditioned matrices of high orders.

133 interval inversion matrices were computed in interval arithmetic. We can expect that the more a matrix will be ill-conditioned the higher will be the condition number $V(A)$.

All computations were done using floating-point arithmetic with a 36-binary-bit mantissa.

Table 1 shows the correlation coefficients among $V(A)$ and measures of errors of the inversion matrix. The measures $\bar{e}^{(1)}$, $\bar{e}^{(2)}$, $\bar{e}^{(3)}$, $\bar{e}^{(4)}$ were computed using ordinary arithmetic. Inversion matrix B was computed by the Jordan inversion method (diagonalization).

Table 1

	$\bar{e}^{(1)}$	$\bar{e}^{(2)}$	$\bar{e}^{(3)}$	$\bar{e}^{(4)}$
$V(A)$	0.643	0.731	0.877	0.883

The notation is as follows:

$A=(a_{ij})$ test matrix, $B=(b_{ij})$ computed inversion matrix, $C=(c_{ij})$ explicitly given inversion matrix.

1. the absolute error $e_{ij}^{(1)}=|b_{ij}-c_{ij}|$.
2. the relative error $e_{ij}^{(2)}=\frac{|b_{ij}-c_{ij}|}{c_{ij}}$.
   If $c_{ij}=0$ then $e_{ij}^{(2)}=0$ was used.
3. the error $e_{ij}^{(3)}=|i_{ij}-d_{ij}|$,
   where $I=(i_{ij})$ is the identity matrix, $D=(d_{ij})=A\,A^{-1}$.
4. the error $e_{ij}^{(4)}=|i_{ij}-f_{ij}|$,
   where $F=(f_{ij})=A^{-1}A$.

The formula used to calculate the average error is:

$$\bar{e}^{(k)}=\frac{\sum\limits_i\sum\limits_j e_{ij}^{(k)}}{n^2}, \quad \text{where} \quad k=1,2,3,4, \quad i,j=1,2,\ldots,n.$$

Strong correlation was proved for $\bar{e}^{(3)}$ and $\bar{e}^{(4)}$. It is known that these values need not reliably indicate the accuracy of computed inversion matrices. See [4] and also [6], p. 119. Strong correlation was also proved for the relative error $\bar{e}^{(2)}$.

Table 2 shows the correlation coefficients among $V(A)$ and these condition numbers:

$$M(A)=n\max_{i,j}|a_{ij}|\max_{i,j}|\alpha_{ij}|, \text{ where } A^{-1}=(\alpha_{ij}).$$

$$N(A)=n^{-1}\Big(\sum_{i,j}a_{ij}^2\Big)^{1/2}\Big(\sum_{i,j}\alpha_{ij}^2\Big)^{1/2},\ i,j=1,2,\ldots,n.$$

$$P(A)=\frac{\max|\lambda_i|}{\min|\lambda_i|}, \text{ where } \lambda_i \text{ are the eigenvalues of } A,\ i=1,2,\ldots,n.$$

$$B(A)=\max_i\sum_j|a_{ij}|\ \max_i\sum_j|\alpha_{ij}|.$$

$$B(A)_{\min}=\pi(|A|\,|A^{-1}|), \text{ where } \pi(X) \text{ is the largest eigenvalue of non-negative matrix } X,\ |A|=(|a_{ij}|).$$

$B(A)_{\min}$ is Bauer's minimum condition number [2]. It gives the limit for ill-conditioning of a matrix which can not be lowered.

Table 2

	$M(A)$	$N(A)$	$P(A)$	$B(A)$	$B(A)_{\min}$
$V(A)$	0.741	0.785	0.768	0.713	0.999

The same set of interval inversion matrices $N=133$ was used for computation of the correlation coefficients.

The correlation coefficients for $M(A)$, $N(A)$, $P(A)$, $B(A)$ are approximately on the same level as the correlation coefficient between $V(A)$ and $\bar{e}^{(2)}$. But the correlation between $V(A)$ and $B(A)_{\min}$ shows very strong dependence. It is nearly a functional dependence and this fact is very interesting.

$B(A)_{\min}$ gives the inherent condition of a matrix. It is the lowest limit which can not be traversed. This condition is a quality of the matrix.

Using the Hansen method we can compute a good approximation of the interval inversion matrix. The condition number $V(A)$ is based on the ratio of average intervals and measures the minimum condition of an interval matrix using interval arithmetic. It is remarkable that the condition numbers $V(A)$ and $B(A)_{\min}$ are constructed from different backgrounds and yet very good accordance exists between them.

Due to the strong dependence mentioned we can find a function for this relation very easily. The least squares method was used.

$$\underset{(0.00176)}{V(A)=1.0145}\,B(A)_{\min}-\underset{(188.1201)}{175.3113}. \tag{6}$$

The standard errors of the regression coefficients are given in parentheses.

Let us denote e_i as the difference between the functional and experimental value. The standard error estimate of e_i is 13 401.58.

The standard error estimate of e_i is relatively high but we must keep in mind that the values of $V(A)$ exist in the range from 10^0 up to 10^8. According to the standard error estimate mentioned I recommend using relation (6) if $B(A)_{\min} > 10^4$. If $B(A)_{\min}$ is lower we can assume well-conditioning of the matrix and expect small interval elements of the interval inversion matrix.

Relation (6) gives the approximation of the condition number $V(A)$ if we know $B(A)_{\min}$. The interpretation is simple and provides the size of interval elements of the interval inversion matrix. This knowledge is important if we know the elements of a matrix with certain accuracy only.

Table 2 shows strong dependence among $V(A)$ and other condition numbers. It is notable that the condition number $V(A)$ is the rough approximation of the condition numbers $P(A)$ or $M(A)$. This can be seen by analyzing the sample mentioned above. According to [5] the condition numbers $P(A)$ and $M(A)$ do not differ too much.

4. Conclusion

It was shown that there exists strong dependence among $V(A)$ and the absolute and the relative error of the inversion matrix. Knowing $V(A)$ we can estimate sizes of interval elements of the interval inversion matrix with respect to the interval elements of the original interval matrix. The condition number $V(A)$ measures the minimum condition of the matrix with high reliability. If we do not know the interval inversion matrix we can estimate $V(A)$ using relation [6].

References

[1] Moore, R. E.: Intervallanalyse, p. 18. München-Wien: R. Oldenbourg 1969.
[2] Bauer, F. L.: Optimally Scaled Matrices. Numerische Mathematik *1963*, 5.
[3] Hansen, E.: Interval Arithmetic in Matrix Computations Part I. Jour. SIAM, Num. Anal. *B 2*, No. 2 (1965).
[4] Porzak, J. P., Denman, H. H.: Evaluation of Error Measures, Condition Numbers and Error Bounds for Certain Matrices. The Journal of the Industrial Mathematics Society *16*, Part 1 (1966).
[5] Zlámal, J.: O podmíněnosti matic (On Condition of Matrices). Doctoral Dissertation, Faculty of Sciences, University of Brno, Czechoslovakia, 1974.
[6] Wilkinson, J. H.: Rundungsfehler. Berlin-Heidelberg-New York: Springer 1969.

Dr. J. Zlámal
Computer Centre
NHKG n. p.
70700 Ostrava-Kunčice
Czechoslovakia