View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Journal of Computational and Applied Mathematics 23 (1988) 117-120
North-Holland

The 2-norm of random matrices *

Per Christian HANSEN
Copenhagen University Observatory, Øster Voldgade 3, DK-1350 København K, Denmark

Received 16 November 1987
Revised 8 February 1988

Abstract: Numerical experiments show that it is possible to derive simple estimates for the expected 2-norm of random matrices A with elements from a normal distribution with zero mean and standard deviation σ, and from a Poisson distribution with mean value λ. These estimates are $\sigma\sqrt{\max(m, n)} < E\{\|A\|_2\} < 2\sigma\sqrt{\max(m, n)}$ and $E\{\|A\|_2\} \approx \lambda\sqrt{mn}$, respectively, where m and n are the dimensions of A.

Keywords: Random matrices, numerical rank decision.

1. Introduction

In a number of technical applications of linear algebra, one is given a matrix whose elements are due to unbiased and uncorrelated absolute or relative errors, usually arising from measurements. It is convenient to consider the perturbed matrix as a sum of a perfect, undisturbed matrix and an error matrix solely consisting of the errors. The 2-norm of the error matrix is of great importance in numerical rank decisions such as in Cholesky factorizations, in Rank Revealing Q-R Factorizations, and via the Singular Value Decomposition, where this norm is used as a threshold for distinction between numerically zero and nonzero quantities, cf. [1,2,3]. The purpose of the present investigation is to derive simple approximations to this norm in order to yield a simple conversion from the measurement errors to the corresponding matrix threshold.

2. Elements from a normal distribution

Consider first the case with unbiased absolute errors and assume that the standard deviations of all matrix elements are identical. This corresponds to an error matrix $A \in \mathbb{R}^{m\times n}$ with elements from a normal distribution with zero mean and standard deviation σ: $a_{ij} \in N(0, \sigma^2)$. It is easy to derive naive upper and lower bounds for the expected value of the 2-norm of this A:

$$E\{\|A\|_2\} \leqslant E\{\|A\|_F\} = \left[\sum E\{a_{ij}^2\}\right]^{1/2} = \left[\sum \sigma^2\right]^{1/2} = \sigma\sqrt{mn}, \tag{1a}$$

$$E\{\|A\|_2\} \geqslant E\{\max(|a_{ij}|)\} = \sigma \tag{1b}$$

* This project was supported by the Danish Space Board.

0377-0427/88/$3.50 © 1988, Elsevier Science Publishers B.V. (North-Holland)

where $E\{\cdot\}$ denotes the expected value, and Σ denotes summation over all i and j. Experience shows that especially the upper bound is much too pessimistic. This is because all the singular values of such a matrix decay only rather slowly, such that $\|A\|_2$ (the largest singular value of A) is far from $\|A\|_F$ (which is the RMS of all the singular values).

To derive more realistic bounds on $E\{\|A\|_2\}$, let $\boldsymbol{r}_i^T$ and $\boldsymbol{a}_j$ denote the rows and columns of A, respectively, and introduce the maximum row/column norm

$$\mu \equiv \max\left[\|\boldsymbol{r}_i\|_2, \|\boldsymbol{a}_j\|_2\right] \tag{2}$$

which is related to m and n by

$$E\{\mu\} \geqslant \max\left[E\{\|\boldsymbol{r}_i\|_2\}, E\{\|\boldsymbol{a}_j\|_2\}\right] = \sigma\sqrt{\max(m, n)}\,. \tag{3}$$

Due to the definition of $\|A\|_2$ it follows that $\mu \leqslant \|A\|_2 \leqslant \sqrt{\max(m, n)}\,\mu$ such that

$$E\{\mu\} \leqslant E\{\|A\|_2\} \leqslant \sqrt{\max(m, n)}\,E\{\mu\}. \tag{4}$$

These bounds are better than those in (1); but they suffer from two inconveniencies; the upper bound is still much too crude, and it is not easy to estimate $E\{\mu\}$ from σ, m, and n.

To derive a simpler lower bound in (4), simply apply (3) to get $\sigma\sqrt{\max(m, n)} \leqslant E\{\|A\|_2\}$. Concerning the upper bound, experience shows that the factor $\sqrt{\max(m, n)}$ should be replaced

Table 1

The results with $a_{ij} \in N(0, \sigma^2)$ with $\sigma = 1$. Each entry contains: First line: mean value $E\{\|A\|_2\}$ and, in ( ), its standard deviation. Second line: the quantity $\sigma\sqrt{\max(m, n)}$. Third line: the quantity $E\{\|A\|_2\}/(\sigma\sqrt{\max(m, n)})$

m	n					
	4	8	16	32	64	128
4	3.1 (0.5) 2.0 1.57	4.1(0.6) 2.8 1.44	5.2 (0.5) 4.0 1.30	7.0 (0.5) 5.7 1.24	9.3 (0.5) 8.0 1.17	12.6 (0.6) 11.3 1.12
8	4.1 (0.5) 2.8 1.45	4.9 (0.6) 2.8 1.75	6.1 (0.5) 4.0 1.53	7.9 (0.6) 5.7 1.40	10.3 (0.5) 8.0 1.28	13.5 (0.5) 11.3 1.20
16	5.2 (0.5) 4.0 1.30	6.1 (0.5) 4.0 1.53	7.3 (0.5) 4.0 1.83	9.1 (0.5) 5.7 1.62	11.5 (0.4) 8.0 1.43	14.7 (0.4) 11.3 1.30
32	6.9 (0.5) 5.7 1.22	7.8 (0.5) 5.7 1.38	9.1 (0.5) 5.7 1.60	10.9 (0.5) 5.7 1.92	13.2 (0.4) 8.0 1.65	16.5 (0.4) 11.3 1.46
64	9.3 (0.5) 8.0 1.16	10.3 (0.5) 8.0 1.28	11.4 (0.5) 8.0 1.43	13.2 (0.4) 8.0 1.65	15.5 (0.4) 8.0 1.94	18.8 (0.4) 11.3 1.66
128	12.7 (0.5) 11.2 1.12	13.5 (0.5) 11.2 1.20	14.8 (0.4) 11.2 1.31	16.5 (0.4) 11.2 1.46	19.0 (0.4) 11.2 1.68	22.2 (0.4) 11.3 1.96

by a much 'milder' function of $\max(m, n)$ and that $E\{\mu\}$ can easily be replaced by its lower bound $\sigma\sqrt{\max(m, n)}$. Equation (4) then becomes:

$$\sigma\sqrt{\max(m, n)} \leqslant E\{\|A\|_2\} \leqslant f(\max(m, n)) \cdot \sigma\sqrt{\max(m, n)} \tag{5}$$

where f is an appropriate simple function which has to be determined.

The bounds in (5) are investigated numerically as follows: random matrices as described above with $\sigma = 1$ were generated in batches of 100, their 2-norm $\|A\|_2$ were computed, and the mean value and standard deviation of $\|A\|_2$ were calculated for each batch and compared with $\sigma\sqrt{\max(m, n)}$. The results are shown in Table 1: the upper line of each entry shows for each batch $E\{\|A\|_2\}$ and its standard deviation, the middle line shows $\sigma\sqrt{\max(m, n)}$, and the bottom line shows the quantity $E\{\|A\|_2\}/(\sigma\sqrt{\max(m, n)})$ which gives the upper and lower bounds for the function f in (5). The table confirms that $\sigma\sqrt{\max(m, n)}$ is indeed a not too pessimistic lower bound on $E\{\|A\|_2\}$. The table also shows that the function f has the asymptotic value 2 for $m = n \to \infty$ and that f always lies between 1 and 2. Similar experiments with $\sigma \neq 1$ gave exactly the same results. Hence, one can use the value 2 for f in (5) to get the following simple bounds for $E\{\|A\|_2\}$:

$$\sigma\sqrt{\max(m, n)} < E\{\|A\|_2\} < 2\sigma\sqrt{\max(m, n)}. \tag{6}$$

This gives the required simple relationship between the standard deviation σ of the matrix elements a_{ij} and the 2-norm of A.

In case the absolute errors have different standard deviations, and in the case of relative errors, it is recommended to apply an initial row and column scaling to A, before further numerical processing, such that as far as possible the standard deviations of the elements of the scaled matrix are all of the same magnitude [3]. The above analysis therefore also applies to these situations.

3. Elements from a Poisson distribution

Consider now the case where all the elements of A belong to a Poisson distribution with parameter $\lambda: a_{ij} \in P(\lambda)$. In this case, the matrix A has non-negative elements of mean value value λ and standard deviation $\sqrt{\lambda}$. Whenever $\lambda > 1$ this means that the first singular value of A, which is identical to $\|A\|_2$, is much larger than the remaining singular values of A. Hence, if $V\{\cdot\}$ denotes the variance, one gets the following simple estimate for $\|A\|_2$:

$$E\{\|A\|_2\} \approx E\{\|A\|_F\} = \left[\sum E\{a_{ij}^2\}\right]^{1/2} = \left[\sum\left[V\{a_{ij}\} + E\{a_{ij}\}^2\right]\right]^{1/2}$$
$$= \left[(\lambda + \lambda^2)mn\right]^{1/2} \approx \lambda\sqrt{mn}. \tag{7}$$

Results from two numerical experiments with $\lambda = 1$ and $\lambda = 5$ are shown in Table 2. The upper two lines show the mean value $E\{\|A\|_2\}$ for batches of 100 matrices generated with $\lambda = 1$ and $\lambda = 5$, respectively, and the bottom two lines show the corresponding values of the ratio $E\{\|A\|_2\}/(\lambda\sqrt{mn})$. It is seen that even for the case $\lambda = 1$, this ratio is never greater than 1.2,

Table 2
The results with $a_{ij} \in P(\lambda)$. Each entry contains: First line: $E\{\|A\|_2\}$ for $\lambda = 1$. Second line: $E\{\|A\|_2\}$ for $\lambda = 5$. Third line: the quantity $E\{\|A\|_2\}/(\lambda\sqrt{mn})$ for $\lambda = 1$. Fourth line: the quantity $E\{\|A\|_2\}/(\lambda\sqrt{mn})$ for $\lambda = 5$.

m	n = 4	8	16	32	64	128
4	4.7	6.4	9.1	12.9	18.0	25.2
	100.7	142.4	200.6	283.5	407.7	568.9
	1.17	1.14	1.14	1.14	1.12	1.12
	1.01	1.01	1.00	1.00	1.00	1.01
8	6.5	8.8	12.5	17.2	24.3	34.1
	141.7	200.9	284.0	400.9	566.6	801.5
	1.15	1.10	1.10	1.07	1.07	1.06
	1.00	1.00	1.00	1.00	1.00	1.00
16	9.3	12.2	17.1	23.6	33.3	46.8
	201.4	284.5	401.6	566.8	800.1	1132.9
	1.16	1.08	1.07	1.05	1.04	1.03
	1.01	1.01	1.00	1.00	1.00	1.00
32	12.8	17.3	23.6	33.0	46.3	65.2
	284.1	401.0	566.1	800.5	1132.3	1601.2
	1.13	1.08	1.04	1.03	1.02	1.02
	1.00	1.00	1.00	1.00	1.00	1.00
64	18.1	23.9	33.5	46.2	65.0	91.5
	401.1	567.1	801.1	1132.2	1601.4	2263.5
	1.13	1.06	1.05	1.02	1.02	1.01
	1.00	1.00	1.00	1.00	1.00	1.00
128	25.5	34.1	46.7	65.4	91.5	129.1
	568.6	802.4	1133.2	1601.6	2263.5	3200.4
	1.13	1.07	1.03	1.02	1.01	1.01
	1.00	1.00	1.00	1.00	1.00	1.00

and for large values of λ it is even closer to 1, without ever getting smaller than 1. Hence, it can be concluded that the quantity $\lambda\sqrt{mn}$ gives a good estimate of $\|A\|_2$, even for λ as small as 1.

4. Conclusion

Our experiments have shown that the singular value spectra of random matrices A with elements from normal and Poisson distributions behave quite differently, such that the estimates for $\|A\|_2$ become rather different. In both cases, however, simple estimates are obtained, giving a very simple relationship between the statistical quantity σ or λ and the matrix norm $\|A\|_2$.

References

[1] T.F. Chan, Rank revealing QR-factorizations, *Linear Alg. Appl.* **88/89** (1987) 67–82.
[2] P.C. Hansen, Detection of near-singularity in Cholesky and LDL^T factorizations, *J. Comput. Appl. Math.* **19** (1987) 293–299.
[3] G.W. Stewart, Rank degeneracy, *SIAM J. Sci. Stat. Comput.* **5** (1984) 403–413.