Journal of Computational and Applied Mathematics 30 (1990) 165–171
North-Holland

Eigenvalues and singular values of certain random matrices

Alan L. ANDREW
Mathematics Department, La Trobe University, Bundoora, Victoria 3083, Australia

Received 28 February 1989
Revised 8 September 1989

Abstract: This paper considers the eigenvalues and singular values of certain matrix-valued random variables whose expected value is a nonnull-matrix of low rank. Implications for the choice of test matrices for numerical algorithms are discussed and some other applications mentioned.

Keywords: Random matrices, eigenvalues, singular values, low rank perturbations, test matrices.

1. Introduction

Matrices for testing numerical algorithms are often generated randomly. The speed of convergence of some iterative methods depends on the ratio of some of the largest eigenvalues of the matrices involved, and the condition of linear systems depends on the ratio of the largest to the smallest of the singular values of the coefficient matrix. Typical values of these ratios for randomly generated matrices will depend on the probability distribution from which the elements are selected. It is clearly desirable that these probability distributions be chosen so that the spread of values of these important ratios should be reasonably representative of that occurring in the matrices for which the algorithms are likely to be used. Yet very little attention seems to be given to this selection in the literature and most authors do not specify the distribution they are using. This paper draws attention to some results concerning the eigenvalues and singular values of certain random matrices whose expected value is not zero. Particular attention is given to a random matrix generator used in the deservedly popular software package MATLAB [9].

The default version of the RAND function of MATLAB generates random matrices with elements independently selected from a uniform distribution on $(0, 1)$. When specifically requested, RAND can also generate normally distributed random matrices with zero mean. However since the default version is likely to be the first choice of some users, an important special property of the (positive) matrices produced in this way is emphasized here. This is that, typically, the dominant (Perron) eigenvalue of these matrices is much larger than the magnitudes of the others, especially for large matrices. Many algorithms perform atypically for such matrices. This fact is discussed briefly in [10] (which does not consider singular values).

Motivated by the importance of such matrices in the numerical solution of ill-posed least squares problems [7], Hansen [6] estimated numerically the expected value of the 2-norm (maximum singular value) of two types of randomly generated matrix, the first with elements from a normal distribution with zero mean (the Gaussian case) and the second with (nonnegative) elements from a Poisson distribution. He observed that this norm increased much more rapidly with the dimensions of the matrix in the latter case and remarked that this was because in this case the largest singular value was much larger than the others. The reason for this dominance of the largest singular value was not mentioned in [6], but essentially it occurs because in this case the expected value of the matrix has rank one and a large nonzero eigenvalue. Random matrices of this type are the subject of our Section 2, while Section 3 considers some random matrices whose expected value has rank greater than one. A fuller investigation of the Gaussian case is given in [3].

The eigenvalues of large symmetric random matrices are also important in quantum mechanics and are the subject of a substantial literature [2]. In most of this, elements are considered to be independent random variables from a normal distribution with zero mean [12], but the case of nonzero mean has also been studied [8]. Other work on eigenvalues of symmetric random matrices has been motivated by applications to graph theory [1]. The role of random matrices in error analysis is also considered in [4]. However there has not been much interaction between those working on different applications.

2. Examples with rank(E(A)) = 1

Let the singular values of the $m \times n$ matrix X be $s_1(X) \geqslant \cdots \geqslant s_n(X)$ and if $m = n$, let the eigenvalues $\lambda_i(X)$ of X be labelled so that $|\lambda_1(X)| \geqslant \cdots \geqslant |\lambda_n(X)|$. Let $\mathbf{A}_{mn}$ denote an $m \times n$ matrix-valued random variable and define $A_{mn} := \mathbf{A}_{mn} - E(\mathbf{A}_{mn})$, where $E(\mathbf{A}_{mn})$ is the expected value of $\mathbf{A}_{mn}$, so that $E(A_{mn}) = 0$. In the important case in which the elements of $\mathbf{A}_{mn}$ are i.i.d. random variables (independent random variables from the same probability distribution) with mean μ and standard deviation σ,

$$E(\mathbf{A}_{mn}) = \mu J_{mn}, \tag{1}$$

where J_{mn} is the $m \times n$ matrix whose elements are all ones. Except where explicitly stated, (1) will be assumed throughout this section. Clearly $J_{mn}^{\mathrm{T}} J_{mn} = m J_{nn}$. The only nonzero singular value of J_{mn} is $(mn)^{1/2}$. Since singular values are relatively insensitive to small changes in the matrix elements, it follows that, when σ is not too large compared with $|\mu|$, $s_1(\mathbf{A}_{mn})$ is likely to be close to $|\mu|(mn)^{1/2}$, and the other singular values much smaller, at least when mn is large. Indeed, since J_{mn} has rank one, it follows from [11, Theorem 1] that, for $i = 2, \ldots, n$,

$$0 \leqslant s_i(\mathbf{A}_{nn}) \leqslant s_{i-1}(A_{nn}) \tag{2}$$

and

$$0 \leqslant s_i(A_{nn}) \leqslant s_{i-1}(\mathbf{A}_{nn}). \tag{3}$$

Most of the literature on random matrices concerns matrices whose expected value is zero. This paper shows how results for such matrices may be used to obtain results for more general random matrices, and in doing so it draws attention to special properties of those whose expected value is nonzero. If the $E(s_i(A_{mn}))$ can be estimated for a given distribution, much

information about $E(s_i(\mathbf{A}_{mn}))$ (and eigenvalues if $m = n$) for arbitrary μ can be gained from (2), (3) and the well-known results

$$s_1^2(\mathbf{A}_{mn}) \leqslant \sum_{i=1}^{n} s_i^2(\mathbf{A}_{mn}) = \|\mathbf{A}_{mn}\|_F^2 := \operatorname{trace}(\mathbf{A}_{mn}^T \mathbf{A}_{mn}), \tag{4}$$

$$\sum_{i=1}^{n} (\lambda_i(\mathbf{A}_{nn}) - \lambda_i(A_{nn})) = \mu n, \tag{5}$$

and the Weyl inequalities [13]

$$\prod_{j=1}^{i} |\lambda_j(\mathbf{A}_{nn})| \leqslant \prod_{j=1}^{i} s_j(\mathbf{A}_{nn}), \quad i = 1, \ldots, n, \tag{6}$$

with equality in (6) when $i = n$. Expanding the right-hand side of (4) shows that its expected value is $mn(\mu^2 + \sigma^2)$. Hence

$$E\left(\sum_{i=1}^{n} (s_i^2(\mathbf{A}_{mn}) - s_i^2(A_{mn})) \right) = mn\mu^2 \tag{7}$$

and since, by the AM–GM inequality, $(\int_a^b f(x)\,dx)^2 \leqslant (b-a)\int_a^b f^2(x)\,dx$ for all $a \leqslant b$ and real-valued functions $f \in L_2[a,b]$, it also follows that

$$E(s_1(\mathbf{A}_{mn})) \leqslant E(\|\mathbf{A}_{mn}\|_F) \leqslant (mn(\mu^2 + \sigma^2))^{1/2}. \tag{8}$$

As $\|\mathbf{A}_{mn}\|_F^2$ is the sum of mn nonnegative i.i.d. random variables, its standard deviation tends to be low compared with its mean, so that $E(\|\mathbf{A}_{mn}\|_F)$ is very close to $(mn(\mu^2 + \sigma^2))^{1/2}$. For individual matrices, $\|\mathbf{A}_{mn}\|_F$ is often larger than this upper bound for its expected value. In particular this is true for the uniform distribution (see Table 1).

By standard perturbation theory [14, pp.87–88], the mild hypothesis

$$2s_1(A_{nn}) < |\mu n| \tag{9}$$

implies that

$$|\lambda_1(\mathbf{A}_{nn}) - \mu n| \leqslant s_1(A_{nn}) \tag{10}$$

Table 1
Results with elements uniformly distributed on $(-\frac{1}{2}, \frac{1}{2})$; sample size k; (the subscripts are omitted from B_{mn} in the headings for simplicity)

m	n	k	Mean value			Standard deviation		
			$s_1(B)$	$s_n(B)$	$\|B\|_F$	$s_1(B)$	$s_n(B)$	$\|B\|_F$
5	5	200	1.02	0.09	1.43	0.11	0.07	0.14
10	10	165	1.58	0.06	2.88	0.13	0.04	0.15
20	5	60	1.69	0.83	2.89	0.13	0.13	0.14
20	20	160	2.37	0.04	5.76	0.12	0.03	0.13
40	10	60	2.53	1.04	5.75	0.13	0.09	0.14
40	40	160	3.48	0.03	11.55	0.11	0.03	0.12
80	20	40	3.68	1.39	11.55	0.12	0.07	0.14
80	80	60	5.04	0.02	23.09	0.10	0.02	0.13

and

$$|\lambda_i(\mathbf{A}_{nn})| \leqslant s_1(A_{nn}), \quad i = 2, \ldots, n. \tag{11}$$

Numerical results indicate that, when $|\mu n|$ is large compared with σ, $\lambda_1(\mathbf{A}_{nn})$ is in fact considerably closer to μn than required by (10) and $|\lambda_1(\mathbf{A}_{nn})|$ is also very close to $s_1(\mathbf{A}_{nn})$. Note that, by (2) and (7),

$$E\big(s_1^2(\mathbf{A}_{nn})\big) \geqslant n^2\mu^2 + E\big(s_n^2(A_{nn})\big) \geqslant n^2\mu^2. \tag{12}$$

Also, when $\mathbf{A}_{nn}$ is Hermitian,

$$|\lambda_i(\mathbf{A}_{nn})| = s_i(\mathbf{A}_{nn}), \quad i = 1, \ldots, n, \tag{13}$$

and, by the minimax theorem [14], the eigenvalues of $\mathbf{A}_{nn}$ and A_{nn} interlace, with

$$\big(\lambda_i(\mathbf{A}_{nn}) - \lambda_i(A_{nn})\big)\mu \geqslant 0, \quad i = 1, \ldots, n. \tag{14}$$

In contrast with $\mathbf{A}_{mn}$ for $\mu \neq 0$, the median of the nonzero singular values of A_{mn} is often of the same order of magnitude as the largest [12]. In this case, for large m and n, $s_1(A_{mn})$ will be much smaller than $\|A_{mn}\|_F$. Since usually $\min(m, n)$ of the $s_i(A_{mn})$ are strictly positive and

$$E\big(\|A_{p^2n,n}\|_F\big) = E\big(\|A_{pn,pn}\|_F\big) \tag{15}$$

for $p = 1, 2, 3, \ldots$, we may expect that

$$E\big(s_1(A_{pn,pn})\big) \leqslant E\big(s_1(A_{p^2n,n})\big) \leqslant pE\big(s_1(A_{pn,pn})\big). \tag{16}$$

Although (16), unlike all other numbered inequalities in this paper, is purely heuristic, the above argument gives a particularly simple explanation of the observed behaviour of singular values of nonsquare matrices. In practice, since $s_n(A_{p^2n,n})$ is generally much larger than the smallest singular values of $A_{pn,pn}$, the upper bound in (16) becomes less sharp as p increases. A stronger asymptotic result for large n has been proved by Geman [3,5]. For a class of distributions which includes the normal distribution studied in [6], his result implies that $E(s_1(A_{m,n})) \sim (m^{1/2} + n^{1/2})\sigma$ as $m, n \to \infty$. The bounds $\sigma(\max(m,n))^{1/2} \leqslant E(s_1(A_{m,n})) \leqslant 2\sigma(\max(m,n))^{1/2}$ deduced from numerical results (but not formally proved) in [6] are clearly compatible with this.

Table 1 shows estimates of the expected value and standard deviation of the maximum and minimum singular values, $s_1(B_{mn})$ and $s_n(B_{mn})$, and Frobenius norm $\|B_{mn}\|_F$ of a random $m \times n$ matrix B_{mn} with elements i.i.d. random variables from a uniform distribution on $(-\frac{1}{2}, \frac{1}{2})$, based on a sample of k such matrices. The results shown for normal and Poisson distributions in [6] include examples with $n > m$ as well as $m > n$, but, since B_{mn} and B_{mn}^T have the same nonzero singular values, the only useful information given by showing these cases separately is the indication of sampling error given by the departure from symmetry in the tables of [6]. Since the standard deviations already give this information, Table 1 shows only the case $m \geqslant n$. Our results indicate that $E(\|B_{mn}\|_F)$ is very close to the upper bound of (8) which in this case has the value

$$\left(mn\int_{-1/2}^{1/2} x^2\,dx\right)^{1/2} = \left(\frac{mn}{12}\right)^{1/2}.$$

As the moments of this uniform distribution are bounded and its variance is $\frac{1}{12}$, Geman's theorem [5] shows that, for all p, $n^{-1/2}s_1(B_{p^2n,n}) \to (1+p)/\sqrt{12}$ almost surely as $n \to \infty$. Table

1 suggests that $n^{-1/2}E(s_1(B_{p^2n,n}))$ increases monotonically to this limit and $E(s_1(B_{2p^2n,2n}))/E(s_1(B_{p^2n,n}))$ decreases monotonically.

The matrices generated by the default version of the RAND command of MATLAB are of the form $B_{mn} + \frac{1}{2}J_{mn}$, their elements being i.i.d. random variables from a uniform distribution on $(0,1)$. The previous discussion shows that the maximum singular value of these matrices will generally be close to $\frac{1}{2}(mn)^{1/2}$, the others being much smaller, and that if $m = n$ the (necessarily real) dominant eigenvalue will be close to $\frac{1}{2}n$ and the others much smaller. Our numerical results suggest a number of other properties.

For many of the matrices B_{mn} used in compiling Table 1, and for some others, we computed the singular values of $B_{mn} + \frac{1}{2}J_{mn}$ and, when $m = n$, the eigenvalues of B_{mn} and $B_{mn} + \frac{1}{2}J_{mn}$. Our results suggest that the ratio $E(|\lambda_1(B_{nn})|)/E(s_1(B_{nn}))$ falls steadily as n increases (and is just over 0.5 for $n = 80$) but that $E(\lambda_1(B_{nn} + \frac{1}{2}J_{nn}))/E(s_1(B_{nn} + \frac{1}{2}J_{nn})) \to 1$ as $n \to \infty$ (increasing from just over 0.9 for $n = 2$ to over 0.99 for $n = 40$). Our results also suggest that in this case $E(|\lambda_2(B_{nn} + \frac{1}{2}J_{nn})|) < E(|\lambda_1(B_{nn})|)$, though there are (necessarily nonsymmetric) matrices B_{nn} for which $|\lambda_2(B_{nn} + \frac{1}{2}J_{nn})| > |\lambda_1(B_{nn})|$. Although (2), (3) give no indication of the relative sizes of $s_n(B_{nn})$ and $s_n(B_{nn} + \frac{1}{2}J_{nn})$, our numerical results suggest that their expected values (and standard deviations) are about the same. There was no obvious correlation between the two values and each seemed to have about a 50% chance of being the smaller. We also calculated the means of the condition numbers $s_1(B_{nn})/s_n(B_{nn})$ and $s_1(B_{nn} + \frac{1}{2}J_{nn})/s_n(B_{nn} + \frac{1}{2}J_{nn})$ and found their ratio was about the same as the ratio of the means of $s_1(B_{nn})$ and $s_1(B_{nn} + \frac{1}{2}J_{nn})$. The distribution of s_n in the Gaussian case is considered further in [3].

While this section has been concerned with matrices whose elements are i.i.d. random variables with nonzero mean, results are readily generalized to random matrices $\mathbf{A}$ with $E(\mathbf{A}) = uv^{\mathrm{T}}$ where u and v are any nonnull real column vectors. The solitary nonzero singular value of uv^{T} is $\|u\|_2 \|v\|_2$ and, if the matrix is square, $u^{\mathrm{T}}v$ is an eigenvalue, the others being zero.

3. Examples with rank($E(\mathbf{A})$) > 1

It would have been of interest to test the algorithms of [10] on some matrices with several relatively large eigenvalues. Changes in a few diagonal elements of the random matrices discussed in Section 2 produce matrices with this property. Here we determine $\lambda_i(E(\mathbf{A}_{nn}))$ for such matrices. Then $E(\lambda_i(\mathbf{A}_{nn}))$ and $E(s_i(\mathbf{A}_{nn}))$ may be estimated by methods similar to those used in Section 2.

Let $\{\alpha_i\}$ be a sequence of nonzero real numbers and let $M_n(r) := J_{nn} + \mathrm{diag}(\alpha_1, \ldots, \alpha_r, 0, \ldots, 0)$ $(r < n)$. Clearly $M_n(r)$ is similar to (and hence has the same eigenvalues as) the matrix obtained by adding the numbers $\alpha_1, \ldots, \alpha_r$ to any distinct diagonal elements of J_{nn}. For every positive integer j and finite set S of at least j numbers, let $\Sigma_{(j)}(S)$ denote the sum of all products of j elements (chosen without replacement) from S, and let $\Sigma_{(0)}(S) := 1$.

Theorem 1. *The $r+1$ nonzero eigenvalues of $M_n(r)$ are the $r+1$ solutions λ of*

$$\sum_{j=0}^{r+1} (-\lambda)^{r+1-j} \sum_{(j)} (n+1-j, \alpha_1, \ldots, \alpha_r) = 0. \tag{17}$$

Proof. Expansion in terms of the ith row shows that

$$\det[M_n(i) - \lambda I] = \det[M_n(i-1) - \lambda I] + \alpha_i \det[M_{n-1}(i-1) - \lambda I],$$

whence, by induction on r,

$$\det[M_n(r) - \lambda I] = \sum_{j=0}^{r+1} (-\lambda)^{n-j} \sum_{(j)} (n+1-j, \alpha_1, \ldots, \alpha_r).$$

The results follows. □

Comparison of (17) with the equation

$$\sum_{j=0}^{r+1} (-\lambda)^{r+1-j} \sum_{(j)} (n, \alpha_1, \ldots, \alpha_r) = 0,$$

whose solutions are $n, \alpha_1, \ldots, \alpha_r$, indicates that, as $n \to \infty$, r solutions will approach $\alpha_1, \ldots, \alpha_r$, the other being close to n. It also follows from the minimax theorem that r of the eigenvalues of $M_n(r)$ are bounded below by $\alpha_1, \ldots, \alpha_r$ respectively, and that, if the α_i are all positive, one eigenvalue is bounded below by n. In the case $r = 1$, (17) becomes

$$\lambda^2 - (n + \alpha_1)\lambda + (n-1)\alpha_1 = 0$$

and, for a given n, the minimum distance between the nonzero eigenvalues occurs when $\alpha_1 = n - 2$, which gives eigenvalues $n - 1 \pm (n-1)^{1/2}$.

Acknowledgements

I thank A. Edelman for drawing my attention to references [1] and [5] and P. Forrester for drawing my attention to reference [2].

References

[1] B. Bollobás, *Random Graphs* (Academic Press, London, 1985).
[2] T.A. Brody, J. Flores, J.B. French, P.A. Mello, A. Pandey and S.S.M. Wong, Random-matrix physics: spectrum and strength fluctuations, *Rev. Modern Phys.* **53** (1981) 385–479.
[3] A. Edelman, Eigenvalues and condition numbers of random matrices, *SIAM J. Matrix Anal. Appl.* **9** (1988) 543–560.
[4] R. Fletcher, Expected conditioning, *IMA J. Numer. Anal.* **5** (1985) 247–273.
[5] S. Geman, A limit theorem for the norm of random matrices, *Ann. Probab.* **8** (1980) 252–261.
[6] P.C. Hansen, The 2-norm of random matrices, *J. Comput. Appl. Math.* **23** (1) (1988) 117–120.
[7] P.C. Hansen, Solution of ill-posed problems by means of truncated SVD, in: R.P. Agarwal, Y.M. Chow and S.J. Wilson, Eds., *Numerical Mathematics, Singapore 1988*, Internat. Ser. Numer. Math. **86** (Birkhäuser, Basel, 1988) 179–192.
[8] R.C. Jones, J.M. Kosterlitz and D.J. Thouless, The eigenvalue spectrum of a large symmetric random matrix with a finite mean, *J. Phys. A* **11** (1978) L45–L48.
[9] C. Moler, MATLAB User's guide, Dept. Comput. Sci., Univ. New Mexico, 1981.
[10] R.C.E. Tan and A.L. Andrew, Computing derivatives of eigenvalues and eigenvectors by simultaneous iteration, *IMA J. Numer. Anal.* **9** (1989) 111–122.

[11] R.C. Thompson, The behaviour of eigenvalues and singular values under perturbations of restricted rank, *Linear Algebra Appl.* **13** (1976) 69–78.
[12] H.F. Trotter, Eigenvalue distributions of large Hermitian matrices; Wigner's semi-circle law and a theorem of Kac, Murdock and Szegö, *Adv. in Math.* **54** (1984) 67–82.
[13] H. Weyl, Inequalities between two kinds of eigenvalues of a linear transformation, *Proc. Nat. Acad. Sci. U.S.A.*, **35** (1949) 408–411.
[14] J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford, 1965).