Canad. Math. Bull. Vol. **27** (3), 1984

# HOW GOOD IS HADAMARD'S INEQUALITY FOR DETERMINANTS?

BY

JOHN D. DIXON

ABSTRACT. Let $A$ be a real $n \times n$ matrix and define the Hadamard ratio $h(A)$ to be the absolute value of $\det A$ divided by the product of the Euclidean norms of the columns of $A$. It is shown that if $A$ is a random variable whose distribution satisfies some simple symmetry properties then the random variable $\log h(A)$ has mean $-\frac{1}{2}n - \frac{1}{4}\log n + 0(1)$ and variance $\frac{1}{2}\log n + 0(1)$. In particular, for each $\varepsilon > 0$, the probability that $h(A)$ lies in the range $[\exp(-\frac{1}{2}n - (\frac{1}{4}+\varepsilon)\log n),\ \exp(-\frac{1}{2}n - (\frac{1}{4}-\varepsilon)\log n)]$ tends to 1 as $n$ tends to $\infty$.

1. **Introduction.** Let $A$ be a real $n \times n$ matrix with columns $a_1, \ldots, a_n$. If the columns are all nonzero, then we define the *Hadamard ratio* $h(A) := |\det A| / \prod_{i=1}^{n} \|a_i\|$ where $\| \ \|$ denotes the Euclidean norm; and if some $a_i = 0$ put $h(A) := 0$. Hadamard's inequality states that $h(A) \le 1$ for all $A$ with equality if and only if the columns of $A$ are mutually orthogonal and no column is 0. The object of this paper is to investigate the distribution of the values of the Hadamard ratio. This investigation has its origin in a comment in [1] which the present author interpreted (or perhaps misinterpreted) to imply that numerical evidence suggests that $h(A)$ is close to 1 for "random" matrices. The theorem below shows that most values of $h(A)^{1/n}$ are close to $e^{-1/2}$.

We must first specify how we shall define a "random" matrix. In fact, our results remain true for quite a wide class of probability distributions. The only conditions on the underlying density of the distribution which we shall need are:

(D1) the density of the distribution at $A$ depends only on the values of $\|a_1\|, \ldots, \|a_n\|$; and

(D2) the probability that $\det A$ is nonzero equals 1.

The following are examples of distributions satisfying (D1) and (D2).

EXAMPLE 1. The uniform distribution over the set of all $n \times n$ matrices $A$ with $\|a_1\| \le \rho_1, \ldots, \|a_n\| \le \rho_n$ for specified constants $\rho_i > 0$.

Research supported in part by NSERC Grant No. A7171.

Received by the editors March 12, 1982 and, in revised form, October 30, 1982.

AMS Subject Classification (1980)

Primary, 15A15; Secondary, 60D05

EXAMPLE 2. The distribution obtained for $n \times n$ matrices $A = [\alpha_{ij}]$ when the $\alpha_{ij}$ are independent random variables from a common normal distribution with mean 0. In this case it is easily seen that the density of the distribution at $A$ depends only on $\sum \alpha_{ij}^2 = \|a_1\|^2 + \cdots + \|a_n\|^2$.

We can now state our main result.

THEOREM. *Let $A$ be a random variable from a distribution satisfying* (D1) *and* (D2). *Let $\mu_n$ and $\sigma_n^2$, respectively, denote the mean and variance of the random variable* $\log h(A)$. *Then*

(i) $\mu_n = -\frac{1}{2}n - \frac{1}{4}\log n + 0(1)$ *and* $\sigma_n^2 = \frac{1}{2}\log n + 0(1)$ *as* $n \to \infty$;

(ii) *for each* $\varepsilon > 0$, *the probability that the inequality*

$$n^{-\frac{1}{4}-\varepsilon}e^{-\frac{1}{2}n} < h(A) < n^{-\frac{1}{4}+\varepsilon}e^{-\frac{1}{2}n}$$

*is satisfied tends to* 1 *as* $n \to \infty$.

REMARK. The mean of $h(A)^2$ has been computed for various underlying probability distributions in the papers [3], [4] and [6] (see also [5]). For $n \times n$ matrices the mean is $n!/n^n$ in each case. (In [6] the expected value of $h(A)^4$ for all matrices with entries 1 or $-1$ is also computed.) In contrast to these results, part (ii) of our Theorem shows much more: not only is the average of $h(A)^2$ near to $n!/n^n$, but nearly all the values of $h(A)^2$ are clustered close to this average.

2. **Basic lemmas.** Let $S_{n-1}$ denote the unit sphere consisting of all $n$-dimensional columns of norm 1.

LEMMA 1. *Let $A$ be a random variable from a distribution satisfying* (D1) *and* (D2). *Then the random variables $u_1 := a_1/\|a_1\|, \ldots, u_n := a_n/\|a_n\|$ are independent and uniformly distributed over $S_{n-1}$.*

REMARK. We use the convention that $a/\|a\| = 0$ if $a = 0$, but (D2) shows that the $a_i$ are nonzero with probability 1.

**Proof.** By (D2) we know that $(u_1, \ldots, u_n) \in (S_{n-1})^n$ with probability 1, and so it is enough to show that the probability density of $(u_1, \ldots, u_n)$ is constant on $(S_{n-1})^n$. Equivalently, we shall show that, for each $\varepsilon > 0$, if $(c_1, \ldots, c_n) \in (S_{n-1})^n$ then $\mathrm{Prob}(\|a_i/\|a_i\| - c_i\| < \varepsilon$ for each $i)$ is independent of $(c_1, \ldots, c_n)$. Suppose $(c_1', \ldots, c_n')$ also lies in $(S_{n-1})^n$. Choose $T_1, \ldots, T_n$ as orthogonal transformations of the $n$-dimensional space such that $T_i c_i = c_i'$ for each $i$. Since $T_i$ preserves distances, $\|a_i/\|a_i\| - c_i\| = \|T_i a_i/\|T_i a_i\| - c_i'\|$. Now (D1) shows that $\mathrm{Prob}(\|a_i/\|a_i\| - c_i'\| < \varepsilon$ for all $i) = \mathrm{Prob}(\|T_i a_i/\|T_i a_i\| - c_i'\| < \varepsilon$ for all $i) = \mathrm{Prob}(\|a_i/\|a_i\| - c_i\| < \varepsilon$ for all $i)$ as required.

LEMMA 2. *Let $x = (\xi_1, \ldots, \xi_n)^T$ be a random variable which is uniformly* *distributed over $S_{n-1}$. For each $r$ with $1 \le r \le n$, the random variable $\eta_{nr} := \sum_{i=1}^{r} \xi_i^2$ has a beta distribution whose density function is*

$$f_{nr}(t) := \frac{\Gamma(\frac{1}{2}n)}{\Gamma(\frac{1}{2}r)\Gamma(\frac{1}{2}(n-r))} t^{\frac{1}{2}r-1}(1-t)^{\frac{1}{2}(n-r)-1} \quad \text{for} \quad 0 \le t \le 1.$$

*Moreover, the random variable $\omega_{nr} := \frac{1}{2}\log \eta_{nr}$ has mean $-\frac{1}{2}\log(n/r) + \frac{1}{2}(n^{-1} - r^{-1}) + 0(r^{-2})$ and variance $\frac{1}{2}r^{-1} + 0(r^{-2})$ with implied constants independent of $n$ and $r$.*

REMARK. A special case of part of this result has been proved in [2].

**Proof.** Let $z = (\zeta_1, \ldots, \zeta_n)^T$ be a random variable whose components $\zeta_i$ are independent random variables with a common normal distribution with mean 0. Then the density of the distribution of $z$ depends only on $\|z\|$, and so $z/\|z\|$ is uniformly distributed over $S_{n-1}$ (compare with proof of Lemma 1). Thus, if we put $\alpha := \sum_{i=1}^{r} \zeta_i^2$ and $\beta := \sum_{i=r+1}^{n} \zeta_i^2$ then $\alpha/(\alpha+\beta)$ has the same distribution as $\eta_{nr}$ has. But $\alpha$ and $\beta$ are random variables with gamma distributions with the parameters $\frac{1}{2}r$ and $\frac{1}{2}(n-r)$, respectively (see [7], p. 172). Therefore $\alpha/(\alpha+\beta)$ (and hence $\eta_{nr}$) has a beta distribution with the parameters $\frac{1}{2}r$, $\frac{1}{2}(n-r)$ (see [7], p. 174), and so $\eta_{nr}$ has the density function $f_{nr}$ given above.

Now consider the characteristic function for $\omega_{nr}$. This is given by

$$\begin{aligned} F(\lambda) &= \int_0^1 \exp(\tfrac{1}{2}\lambda \log t) f_{nr}(t)\, dt \\ &= \frac{\Gamma(\frac{1}{2}n)}{\Gamma(\frac{1}{2}r)\Gamma(\frac{1}{2}(n-r))} \int_0^1 t^{\frac{1}{2}(r+\lambda)-1}(1-t)^{\frac{1}{2}(n-r)-1}\, dt \\ &= \frac{\Gamma(\frac{1}{2}n)\Gamma(\frac{1}{2}(r+\lambda))}{\Gamma(\frac{1}{2}r)\Gamma(\frac{1}{2}(n+\lambda))} \quad \text{for all} \quad \lambda > -r. \end{aligned}$$

Since

$$\Gamma'(t)/\Gamma(t) = -\gamma - t^{-1} + \sum_{k=1}^{\infty} \{k^{-1} - (k+t)^{-1}\} \quad \text{for all} \quad t > 0$$

we conclude that

$$F'(\lambda)/F(\lambda) = \sum_{k=0}^{\infty} \{(n+\lambda+2k)^{-1} - (r+\lambda+2k)^{-1}\}$$

and

$$F''(\lambda)/F(\lambda) - \{F'(\lambda)/F(\lambda)\}^2 = -\sum_{k=0}^{\infty} \{(n+\lambda+2k)^{-2} - (r+\lambda+2k)^{-2}\}.$$

Now Euler's summation formula shows that for $s > 0$ and $j > 0$ we have

$$\sum_{k=0}^{m} (s+2k)^{-j} = \int_0^m (s+2t)^{-j}\, dt + \tfrac{1}{2}\{s^{-j} + (s+2m)^{-j}\} + R_m$$

https://doi.org/10.4153/CMB-1984-039-2 Published online by Cambridge University Press

where

$$R_m = -j\int_0^m (t-[t]-\tfrac{1}{2})(s+2t)^{-j-1}\,dt$$
$$= -j\sum_{k=0}^{2m-1}\int_{\frac{1}{2}k}^{\frac{1}{2}(k+1)} (t-[t]-\tfrac{1}{2})(s+2t)^{-j-1}\,dt.$$

Since $R_m$ is an alternating sum whose terms are decreasing in absolute value, we deduce that

$$0 < R_m < j\int_0^{\frac{1}{2}} \tfrac{1}{2}(s+2t)^{-j-1}\,dt = \tfrac{1}{2}\{s^{-j}-(s+1)^{-j}\}.$$

Thus, since $F(0)=1$, we conclude that the mean $F'(0)$ of $\omega_{nr}$ equals

$$\sum_{k=0}^{\infty}\{(n+2k)^{-1}-(r+2k)^{-1}\} = -\tfrac{1}{2}\log(n/r)+\tfrac{1}{2}(n^{-1}-r^{-1})+0(r^{-2})$$

and similarly the variance $F''(0)-F'(0)^2$ of $\omega_{nr}$ equals $\frac{1}{2}r^{-1}+0(r^{-2})$.

3. **Proof of the theorem.** We first observe that $h(A)$ is equal to the determinant of the matrix with columns $a_1/\|a_1\|, \ldots, a_n/\|a_n\|$. Hence Lemma 1 shows that the distribution of the random variable $h(A)$ is identical to the distribution of $|\det(u_1, \ldots, u_n)|$ where $u_1, \ldots, u_n$ are independent random variables uniformly distributed over $S_{n-1}$. On the other hand, $|\det(u_1, \ldots, u_n)|$ is equal to the volume of the $n$-dimensional parallelopiped with edges $u_1, \ldots, u_n$; and so $|\det(u_1, \ldots, u_n)| = \pi_2\pi_3\cdots\pi_n$ where $\pi_{i+1}$ is the length of the projection of $u_{i+1}$ into the orthogonal complement of the subspace $V_i$ spanned by $u_1, \ldots, u_i$. Suppose now that $\det(u_1, \ldots, u_n)\neq 0$ (this occurs with probability 1). Then each $V_i$ has dimension $i$, and there is an orthogonal transformation $T_i$ of the $n$-dimensional space which maps $V_i$ onto the subspace $W_i$ consisting of all vectors whose first $n-i$ components are 0. Now suppose $u_1, \ldots, u_i$ are fixed. Since $T_i$ is measure-preserving and distance-preserving, $T_iu_{i+1}$ is a random variable which is uniformly distributed over $S_{n-1}$ and the length of its projection into the orthogonal complement of $T_iV_i = W_i$ equals $\pi_{i+1}$. Thus, for fixed $u_1, \ldots, u_i$, the distribution of $\pi_{i+1}^2$ is identical to that of the variable $\eta_{n,n-i}$ defined in Lemma 2, and so $\log \pi_{i+1}$ is distributed identically to $\omega_{n,n-i}$. However, the $u_i$ are independent random variables, and so the distribution of $\log \pi_{i+1}$ does not depend on $u_1, \ldots, u_i$, and the variables $\pi_2, \ldots, \pi_n$ are also independent.

Thus, we have shown that the distribution of $\log h(A)$ is the same as the distribution of the sum $\sum_{i=2}^n \log \pi_i$ of independent random variables where $\log \pi_{i+1}$ is distributed like $\omega_{n,n-i}$. Thus the mean of $\log h(A)$ is the sum of the means of $\omega_{n,r}$ $(r=1, \ldots, n-1)$ and the variance of $\log h(A)$ is the sum of the variances.

Hence Lemma 2 shows that

$$\mu_n = \sum_{r=1}^{n-1} \{-\tfrac{1}{2}\log(n/r) + \tfrac{1}{2}n^{-1} - \tfrac{1}{2}r^{-1} + O(r^{-2})\} = -\tfrac{1}{2}n - \tfrac{1}{4}\log n + O(1)$$

and

$$\sigma_n^2 = \sum_{r=1}^{n-1} \{\tfrac{1}{2}r^{-1} + O(r^{-2})\} = \tfrac{1}{2}\log n + O(1).$$

This proves part (i) of the Theorem. Part (ii) follows immediately using the Chebychev inequality.

ACKNOWLEDGEMENTS. The final form of this paper has benefitted very considerably from suggestions and criticisms of Brian Mortimer, Donald Dawson and, particularly, Paul Beesack, of an earlier attempt.

## REFERENCES

1. S. Cabay and T. P. L. Lam, *Congruence techniques for the exact solution of integer systems of linear equations*, ACM Trans. Math. Software **3** (1977) 386–397.

2. J. D. Dixon, *Estimating extremal eigenvalues and condition numbers of matrices*, SIAM J. Numer. Anal., **20** (1983), 812–814.

3. C. R. Johnson and M. Newman, *How bad is the Hadamard determinantal bound*?, J. Res. Nat. Bureau Stand. **78B** (1974) 167–169.

4. J. S. Shiue, *On a generalization of a theorem of Johnson and Newman*, Soochow J. Math. Natur. Sci. **2** (1976) 57–61.

5. G. Szekeres, *The average value of skew Hadamard matrices, in* Proc. First Austral. Conf. Combinatorial Math. (Newcastle, 1972), pp. 55–59.

6. G. Szekeres and P. Turán, *On an extremal problem in determinant theory*, Mat. természett. Ertes. **56** (1937) 796–804 (Hungarian) (see Zbl. Math. **18** (1938) 387).

7. S. S. Wilks, *Mathematical Statistics*, Wiley, New York, 1962.

DEPARTMENT OF MATHEMATICS AND STATISTICS
CARLETON UNIVERSITY
OTTAWA, CANADA K1S 5B6