How Tight is Hadamard's Bound?

John Abbott and Thom Mulders

CONTENTS

For a real square matrix M, Hadamard's inequality gives an upper bound H for the determinant of M; the bound is sharp if and only if the rows of M are orthogonal. We study how much we can expect that H overshoots the determinant of M, when the rows of M are chosen randomly on the surface of the sphere. This gives an indication of the "wasted effort" in some modular algorithms.

1. INTRODUCTION

Let n be a positive integer and $M \in \mathbb{R}^{n \times n}$ a real n by n matrix. The only bound for the absolute value of the determinant d of M of which we are aware is that due to Hadamard (see [Horn and Johnson 1985], for instance):

$$H = \prod_{i=1}^{n} |v_i| \geq |d|,$$

where v_i is the vector whose coordinates are given by row i of M, and $|v|$ represents the euclidean length of the vector v. We can equally well use the columns of the matrix instead of its rows. Indeed, the smaller of these two possibilities is still a valid upper bound. Empirically we have observed that the row and column bounds rarely differ by a factor greater than 2 for matrices with independent uniformly distributed entries in some fixed range. Nonetheless there do exist cases where the two values differ significantly, such as an identity matrix with all entries in one column replaced by a large integer K. So, in practical applications, since both versions of the bound can be computed cheaply it is worth doing so, and then using the smaller.

For the analysis in this paper we shall assume that just one of the versions of Hadamard's inequality is used, say the "row" version. Hadamard's bound is tight if and only if the rows of the matrix are orthogonal; otherwise it is pessimistic. The relative error

$H/|d|$ is sometimes referred to as the *orthogonality defect* (see [Clarkson 1992], for example). However, to avoid difficulties which arise when $d = 0$ we shall work instead with $|d|/H$. In particular, we shall determine the expectations and variances of the random variables $|d|/H$ and $\log(|d|/H)$ for matrices whose rows are random vectors uniformly distributed on the surface of a sphere. The former random variable takes values in the interval $[0,1]$ while the latter is nonpositive.

The interest in the tightness of Hadamard's bound is stimulated by algorithms for computing exactly the determinant d of an integer matrix. One way is to use Chinese remaindering: compute d modulo several distinct primes, $p_1, p_2, \ldots, p_k$, and then combine these modular images of d [Bareiss 1972]. We obtain the correct answer provided that $\prod_{i \in I} p_i > 2|d|$, but since we do not generally know $|d|$ in advance we use instead the sufficient condition that $\prod_{i \in I} p_i > 2H$. Thus $\log(H/|d|)$ is a measure of the "wasted effort" using this method.

Another way of computing d is given in [Abbott et al. 1999] and works by computing quickly some (probably large) factor D of $|d|$, and then using Chinese remaindering to find the value of d/D. A suitable D arises as the least common denominator of the solution of a linear system $Mx = b$ for a random integer vector b. Once again $\log(H/|d|)$ is related to the "wasted effort", and the same article gave asymptotic upper and lower bounds for its mean. In this paper we improve these bounds and give explicit formulae for both mean and variance of $|d|/H$ and of $\log(|d|/H)$.

Henceforth we shall assume that the matrix dimension n is at least 2, the case $n = 1$ being trivial.

2. DISTRIBUTION

The first point to establish is our assumed distribution on the matrices. We regard the matrix dimension n as being arbitrary but fixed. We shall avoid all matrices containing a row of zeroes because they give $d = H = 0$. Now, the value of $|d|/H$ is unaffected if we scale any row of the matrix by a nonzero value; thus we may assume that each row has euclidean length 1, and consequently that $H = 1$. Such matrices can be viewed as specifying n points, $v_1, \ldots, v_n$, on the surface of the unit

n-sphere. Our assumed distribution is that these points are independently and uniformly distributed on the surface of the sphere. Thus our study becomes that of the random variables $|d|$ and $\log|d|$.

Now, viewing the absolute value of the determinant as the volume of a parallelepiped we see that it can be computed as a product of sines of certain angles. We define $n - 1$ angles as follows: let φ_1 be the angle between v_1 and v_2, let φ_2 be the angle between $\mathrm{span}(v_1, v_2)$ and v_3, and in general let φ_k be the angle between $\mathrm{span}(v_1, \ldots, v_k)$ and v_{k+1} for k up to $n - 1$. It is clear that

$$|d| = \sin \varphi_1 \sin \varphi_2 \cdots \sin \varphi_{n-1}.$$

We focus our attention on φ_k for the moment. We begin by finding the density function for φ_k, and by symmetry we can suppose that $\varphi_k \in [0, \pi/2]$. Let $W_k = \mathrm{span}(v_1, \ldots, v_k)$, a subspace of dimension k. The vector v_{k+1} can be written uniquely as the sum of two components: $v_\parallel$ lying in W_k and $v_\perp$ perpendicular to W_k. Furthermore $|v_\parallel| = \cos \varphi_k$ and $|v_\perp| = \sin \varphi_k$ by definition. Thus all unit vectors forming an angle φ_k with W_k are generated as the sum of a vector lying on the k-sphere of radius $\cos \varphi_k$ inside W_k with another vector lying on the $(n - k)$-sphere of radius $\sin \varphi_k$ inside $W_k^\perp$. Whence we see that the density of φ_k as a function in θ is proportional to $(\sin \theta)^{n-k-1} (\cos \theta)^{k-1}$.

3. ANALYSIS OF $|d|/H$

Since $H = 1$ for our assumed distribution, we need only study $|d|$. We determine first its expectation, then its variance. Thanks to the symmetry of the sphere, the angles $\varphi_1, \ldots, \varphi_{n-1}$ are mutually independent, and so

$$\boldsymbol{E}(|d|) = \prod_{k=1}^{n-1} \boldsymbol{E}(\sin \varphi_k). \qquad (3\text{–}1)$$

Let Γ be the usual gamma function and let B be the beta function, defined by

$$B(p,q) = \int_0^1 x^{p-1} (1 - x)^{q-1} \, dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p + q)} \qquad (3\text{–}2)$$

[Gröbner and Hofreiter 1961, 421.1]. Then, for nonnegative integers r, s, we have

$$\int_0^{\pi/2} (\sin \theta)^r (\cos \theta)^s \, d\theta = \tfrac{1}{2} B\left(\tfrac{r+1}{2}, \tfrac{s+1}{2}\right) \qquad (3\text{–}3)$$

[Gröbner and Hofreiter 1961, 331.21]. (The facts we will need about special functions can also be found in [Abramowitz and Stegun 1972; Gradshteyn and Ryzhik 1980; Luke 1969].)

Using (3–3) and (3–2) to integrate our density function, we get

$$E(\sin \varphi_k) = \frac{\displaystyle\int_0^{\pi/2} (\sin \theta)^{n-k} (\cos \theta)^{k-1} \, d\theta}{\displaystyle\int_0^{\pi/2} (\sin \theta)^{n-k-1} (\cos \theta)^{k-1} \, d\theta}$$

$$= B\left(\tfrac{n-k+1}{2}, \tfrac{k}{2}\right) \big/ B\left(\tfrac{n-k}{2}, \tfrac{k}{2}\right)$$

$$= \frac{\Gamma\left(\tfrac{n}{2}\right) \Gamma\left(\tfrac{n-k+1}{2}\right)}{\Gamma\left(\tfrac{n+1}{2}\right) \Gamma\left(\tfrac{n-k}{2}\right)}.$$

By substituting this in (3–1) and telescoping the product we get

Lemma 3.1. $E(|d|) = \dfrac{\Gamma\left(\tfrac{n}{2}\right)^n}{\Gamma\left(\tfrac{1}{2}\right) \Gamma\left(\tfrac{n+1}{2}\right)^{n-1}}.$

The asymptotic behaviour of $E(|d|)$ can now be derived easily from the equality $\Gamma\left(\tfrac{1}{2}\right) = \sqrt{\pi}$ and from the asymptotic expansion of the logarithm of the gamma function, which is

$$\log \Gamma(x) = \left(x - \tfrac{1}{2}\right) \log x - x + \frac{\log(2\pi)}{2} + \frac{1}{12\,x} + O(x^{-2})$$

for $x > 0$. After suitable simplification we obtain the following lemma.

Lemma 3.2. $E(|d|) = e^{-n/2} \sqrt[4]{4e} \left(1 + O(n^{-1})\right).$

Informally, this means that for large n on average $\det M$ is smaller than Hadamard's bound by roughly a factor of $e^{-n/2} \sqrt[4]{4e}$.

We cannot compute $E(1/|d|)$ since we encounter an unbounded integral:

$$E(1/\sin \varphi_{n-1}) = \frac{\displaystyle\int_0^{\pi/2} \frac{(\cos \theta)^{n-2}}{\sin \theta} \, d\theta}{\displaystyle\int_0^{\pi/2} (\cos \theta)^{n-2} \, d\theta} = \infty.$$

Variance of $|d|$

Further understanding of the behaviour of $|d|$ comes from knowing its variance. We shall use the well-known formula $\operatorname{var} |d| = E(|d|^2) - (E(|d|))^2$ for our computation. We have obtained $E(|d|)$ above, so turn immediately to $E(|d|^2)$.

Again, since the φ_k are mutually independent, we have

$$E(|d|^2) = \prod_{k=1}^{n-1} E\left((\sin \varphi_k)^2\right). \qquad (3\text{–}4)$$

As before we use the density function and equations (3–3) and (3–2), together with the functional relation $\Gamma(x + 1) = x\Gamma(x)$, to get

$$E(\sin^2 \varphi_k) = \frac{\displaystyle\int_0^{\pi/2} (\sin \theta)^{n-k+1} (\cos \theta)^{k-1} d\theta}{\displaystyle\int_0^{\pi/2} (\sin \theta)^{n-k-1} (\cos \theta)^{k-1} d\theta}$$

$$= B\left(\tfrac{n-k+2}{2}, \tfrac{k}{2}\right) \big/ B\left(\tfrac{n-k}{2}, \tfrac{k}{2}\right)$$

$$= \frac{n - k}{n}.$$

Substituting this in (3–4) we get

$$E(|d|^2) = \frac{\Gamma(n)}{n^{n-1}} = \frac{\Gamma(n + 1)}{n^n}.$$

Applying the asymptotic expansion for the logarithm of the gamma function we obtain:

$$E(|d|^2) = e^{-n} \sqrt{2\pi n} \left(1 + O(n^{-1})\right).$$

Combining this with Lemma 3.2 we get the asymptotic behaviour of the variance:

Lemma 3.3. $\operatorname{var} |d| = e^{-n} (\sqrt{2\pi n} - \sqrt{4e}) \left(1 + O(n^{-1})\right).$

Chebyshev's Inequality [Lukacs 1971] relating variance to probable deviation from the mean now tells us that indeed $|d|$ is unlikely to stray too far from the expected value: with a probability of at least 95% we have that $|d|/E(|d|)$ does not exceed $4n^{1/4}$ for large n, though we get no information about how much smaller than its mean $|d|$ might be in multiplicative terms.

4. ANALYSIS OF $\log(|d|/H)$

Once again recall from section 2 that we have $H = 1$ for our assumed distribution, so in fact we need only study $\log |d|$. In [Abbott et al. 1999] bounds for its expectation were obtained using Jensen's inequality [Breiman 1968]. Here we obtain an exact formula as well as the first terms of its asymptotic expansion (using elementary functions). Afterwards we ascertain the variance finding both exact and asymptotic formulae.

The mutual independence of the angles $\varphi_1, \ldots, \varphi_{n-1}$ allows the simplification

$$\boldsymbol{E}(\log|d|) = \sum_{k=1}^{n-1} \boldsymbol{E}(\log \sin \varphi_k). \qquad (4\text{--}1)$$

To write this in closed form we use the logarithmic derivative of the gamma function, denoted by Ψ; we have, for nonnegative integers r, s:

$$\int_0^{\pi/2} (\sin \theta)^r (\cos \theta)^s \log \sin \theta \, d\theta$$
$$= \tfrac{1}{4} B\left(\tfrac{r+1}{2}, \tfrac{s+1}{2}\right) \left(\Psi\left(\tfrac{r+1}{2}\right) - \Psi\left(\tfrac{r+s+2}{2}\right)\right) \quad (4\text{--}2)$$

[Gröbner and Hofreiter 1961, 338.6c].

Much as before, writing the expectation as an integral using the density function for φ_k and then employing (3–3) and (4–2) we get

$$\boldsymbol{E}(\log \sin \varphi_k) = \frac{1}{2} \left(\Psi\left(\frac{n-k}{2}\right) - \Psi\left(\frac{n}{2}\right)\right). \quad (4\text{--}3)$$

We reduce the sum to a single closed form expression:

Lemma 4.1. *The expectation of* $\log|d|$ *equals*

$$\frac{2 - \gamma - 2\log 2}{4} - \frac{n}{2} - \frac{n}{4}\Psi\left(\frac{n}{2}\right) + \frac{n-1}{4}\Psi\left(\frac{n+1}{2}\right).$$

Proof. Substituting (4–3) into (4–1) and inverting the summation order we get

$$\boldsymbol{E}(\log|d|) = \frac{1}{2} \sum_{k=1}^{n-1} \left(\Psi\left(\frac{k}{2}\right) - \Psi\left(\frac{n}{2}\right)\right). \quad (4\text{--}4)$$

Let $f(n)$ denote the expression in the statement of the lemma and $g(n)$ the sum in (4–4). A straightforward computation, using the value for $\Psi(\frac{1}{2}) = -\gamma - 2\log 2$ (where γ is Euler's constant) and the functional equation

$$\Psi\left(\frac{n+2}{2}\right) = \Psi\left(\frac{n}{2}\right) + \frac{2}{n},$$

shows that $f(1) = g(1)$ and that $f(n+1) - f(n) = g(n+1) - g(n)$ for all $n \geq 1$. Whence the lemma follows. $\square$

To get a better idea of the magnitude of $\boldsymbol{E}(\log|d|)$ we use the asymptotic expansion of $\Psi(x)$, which is $\log x - 1/(2x) + O(x^{-2})$ for $x > 0$. After simplification we get:

Lemma 4.2. *The expectation of* $\log|d|$ *is given by*

$$-\frac{n}{2} - \frac{\log n}{4} - \frac{\gamma + \log 2 - 3}{4} + O(n^{-1}).$$

Notice that $\boldsymbol{E}(\log|d|)$ is similar to but not equal to $\log(\boldsymbol{E}|d|)$. We have

$$\log(\boldsymbol{E}(|d|)) - \boldsymbol{E}(\log|d|) = \frac{\log n}{4} + O(1).$$

Of course, that $\log(\boldsymbol{E}(|d|)) \geq \boldsymbol{E}(\log|d|)$ follows qualitatively from Jensen's inequality [Breiman 1968].

Informally, this lemma states that the "wasted effort" mentioned in the introduction is linear in the dimension n of the matrix and is independent of the size of entries of the matrix.

Variance of $\log|d|$

We now compute the variance of $\log|d|$. Yet again we use the mutual independence of the φ_k to simplify the computations: the variance of a sum of independent random variables is just the sum of their individual variances, so

$$\operatorname{var} \log|d| = \sum_{k=1}^{n-1} \left(\boldsymbol{E}((\log \sin \varphi_k)^2) - (\boldsymbol{E}(\log \sin \varphi_k))^2\right). \qquad (4\text{--}5)$$

We express the expectation of $(\log \sin \varphi_k)^2$ as an integral, and then simplify using (3–3) and the formula

$$\int_0^{\pi/2} (\sin \theta)^r (\cos \theta)^s (\log \sin \theta)^2 \, d\theta = \tfrac{1}{8} B\left(\tfrac{r+1}{2}, \tfrac{s+1}{2}\right)$$
$$\times \left(\Psi'\left(\tfrac{r+1}{2}\right) - \Psi'\left(\tfrac{r+s+2}{2}\right) + \left(\Psi\left(\tfrac{r+1}{2}\right) - \Psi\left(\tfrac{r+s+2}{2}\right)\right)^2\right),$$

valid for nonnegative integers r, s (this follows from (3–2) and [Gröbner and Hofreiter 1961, 324.8a]). We obtain

$$\boldsymbol{E}((\log \sin \varphi_k)^2)$$
$$= \tfrac{1}{4} \left(\Psi'\left(\tfrac{n-k}{2}\right) - \Psi'\left(\tfrac{n}{2}\right) + \left(\Psi\left(\tfrac{n-k}{2}\right) - \Psi\left(\tfrac{n}{2}\right)\right)^2\right). \qquad (4\text{--}6)$$

The sum representing the variance of $\log|d|$ can now be telescoped to produce a more compact form for which we will afterwards show the asymptotic behaviour. A closed form is given by the following result:

Lemma 4.3. *The variance of* $\log|d|$ *is given by*

$$\frac{8\gamma + \pi^2}{16} + \frac{\Psi(n)}{2} + \frac{n-1}{8}\Psi'\left(\frac{n+1}{2}\right) - \frac{n}{8}\Psi'\left(\frac{n}{2}\right).$$

Proof. Putting together formulae (4–5), (4–6) and (4–3) and inverting the summation order we obtain

$$\operatorname{var}\log|d| = \frac{1}{4}\sum_{k=1}^{n-1}\left(\Psi'\left(\frac{k}{2}\right) - \Psi'\left(\frac{n}{2}\right)\right) \qquad (4\text{–}7)$$

Let $f(n)$ denote the expression in the statement of the lemma and $g(n)$ the sum in (4–7). A straightforward computation, using the values $\Psi(1) = -\gamma$, $\Psi'(\frac{1}{2}) = \frac{1}{2}\pi^2$ and the functional equations

$$\Psi(n+1) = \Psi(n) + \frac{1}{n} \quad\text{and}\quad \Psi'\left(\frac{n+2}{2}\right) = \Psi'\left(\frac{n}{2}\right) - \frac{4}{n^2},$$

shows that $f(1) = g(1)$ and that $f(n+1) - f(n) = g(n+1) - g(n)$ for all $n \geq 1$. Whence the lemma follows. $\qquad\square$

Directly applying the asymptotic expansions of $\Psi(x)$ (given after the proof of Lemma 4.1) and of $\Psi'(x)$, we get

Lemma 4.4. $\operatorname{var}\log|d| = \frac{1}{16}(8\gamma+\pi^2)+\frac{1}{2}\log n+O(n^{-1})$.

Here again Chebyshev's Inequality shows that for a given probability $\varepsilon > 0$ actual values of $\log|d|$ will lie within an interval of size $O(\sqrt{\log n})$ with probability $1 - \varepsilon$. In contrast with what we could deduce from our analysis of $|d|$, here the interval gives us both upper and lower probable bounds for $|d|$. Furthermore, for sufficiently large n the upper bound obtained here is tighter than the one derived earlier.

5. EXPERIMENTS

We have compared the theoretical results of Section 4 with practice by performing some experiments. For this we have computed the determinant and Hadamard's bound for many matrices. The rows of the matrices were constructed as follows:

1. Determine a random vector on the unit sphere (uniformly distributed);
2. Multiply the vector by 10^6;
3. Round every entry of the vector to an integer.

In this way the rows of the matrices approximate random integer vectors on the sphere of radius 10^6. We wanted the matrices to have integer entries in order to avoid errors in the determinant computation due to numerical instability.

n	N	Predicted		Experimental	
		E	var	mean	var
2	10000	−0.693	0.822	−0.690	0.791
3	10000	−1.307	1.178	−1.297	1.141
10	2000	−5.131	1.979	−5.117	1.916
20	2000	−10.310	2.365	−10.379	2.464

TABLE 1. Comparison of theoretical and experimental results for $\log|d|$ (mean and variance). The values under "Predicted" come from the formulas of Section 4, and those under "Experimental" are the statistics for a sample of N matrices of size $n \times n$.

Table 1 lists the results of these experiments. We see that there is good accord with the theoretical results; The expected value of $\log(|d|/H)$ and the experimental mean differ by less than 1%, and the theoretical and experimental variances differ by less than 5%.

6. CONCLUSIONS

We have derived formulae for the expected values and variances of $|d|/H$ and $\log(|d|/H)$ for a random matrix M, where H is Hadamard's bound for M and d is its determinant. For all these quantities we have also given asymptotic expansions in terms of the matrix dimension n. These results improve the bounds given in [Abbott et al. 1999].

Our results give an idea of the tightness of Hadamard's bound in the average case. Moreover, the small variances, particularly that of $\log(|d|/H)$, indicate that observed values should normally lie close to the average.

The value $\log(H/|d|) = -\log(|d|/H)$ is a measure of the "wasted effort", when computing the determinant of an integer matrix using a Chinese remaindering method. We have seen that this "wasted effort" is on average linear in n with typical deviation of size $O(\sqrt{\log n})$ which is negligible in comparison. Observe also that the "wasted effort" is independent of the size of entries in M.

REFERENCES

[Abbott et al. 1999] J. Abbott, M. Bronstein, and T. Mulders, "Fast deterministic computation of determinants of dense matrices", pp. 197–204 in *Proc. 1999 International Symposium on Symbolic and Algebraic*

Computation (*ISSAC*) (Vancouver, BC), ACM, New York, 1999.

[Abramowitz and Stegun 1972] *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, edited by M. Abramowitz and I. A. Stegun, Dover, New York, 1972. Reprinted 1992.

[Bareiss 1972] E. H. Bareiss, "Computational solutions of matrix problems over an integral domain", *J. Inst. Math. Appl.* **10** (1972), 68–104.

[Breiman 1968] L. Breiman, *Probability*, Addison-Wesley Publishing Company, Reading, Mass., 1968. Reprinted with corrections by SIAM, Philadelphia, 1992.

[Clarkson 1992] K. L. Clarkson, "Safe and effective determinant evaluation", pp. 387–395 in Proc. 33rd Ann. IEEE Symp. Foundations of Comp. Science (Pittsburgh, PA, 1992), IEEE, Los Alamitos, CA 1992, 1992.

[Gradshteyn and Ryzhik 1980] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, Corrected and enlarged ed., Academic Press, New York, 1980.

[Gröbner and Hofreiter 1961] W. Gröbner and N. Hofreiter, *Integraltafel, Teil 2: Bestimmte Integrale*, Springer, Vienna, 1961.

[Horn and Johnson 1985] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1985. Reprinted with corrections 1990.

[Lukacs 1971] E. Lukacs, *Probability and mathematical statistics. An introduction*, Academic Press, New York, 1971.

[Luke 1969] Y. L. Luke, *The special functions and their approximations*, Academic Press, New York, 1969.

John Abbott, Dipartimento di Matematica, Università di Genova, Via Dodecaneso 35, I-16146 Genova, Italy (abbott@dima.unige.it)

Thom Mulders, Institute of Scientific Computing, ETH Zürich, Haldeneggsteig 4, CH-8092 Zürich, Switzerland (thom.mulders@comit.ch)