Probabilistic Lower Bounds on Maximal Determinants of Binary Matrices

Richard P. Brent
Australian National University

Judy-anne Osborn
University of Newcastle, NSW

16 June 2016

Joint work with Warren D. Smith

Probabilistic and Extremal Combinatorics Downunder,
Monash University, 13–17 June 2016

The Hadamard maximal determinant problem

Let A be a $\{\pm 1\}$-matrix of order n,
i.e. an $n \times n$ matrix with entries in $\{-1, +1\}$.

How large can $\det(A)$ be?

Hadamard (1893) partly answered the question by proving an upper bound

$$|\det(A)| \le n^{n/2}$$

that can be attained for infinitely many values of n. Such n are called *Hadamard orders* and the matrices attaining the bound are called *Hadamard matrices*.

Hadamard's bound can be proved by applying the arithmetic-geometric mean inequality to the eigenvalues of $A^T A$.

Some variants of the maxdet problem

- We can ask the same question for $n \times n$ matrices that are allowed to have real entries in $[-1, 1]$. Since the maxima occur at extreme points of $[-1, 1]^n$, the answer is the same as before.
- We can ask the same question for $(n-1) \times (n-1)$ matrices whose entries are in $\{0, 1\}$ or $[0, 1]$. The answer is the same, except for a scaling factor of 2^{n-1}.
- A more general problem is to maximise $\det(A^T A)$, where A is an $m \times n$ matrix with entries in $\{-1, +1\}$, and $m \geq n$. This problem arises in the design of experiments.

The functions $D(n)$ and $R(n)$

Let $D(n)$ be the maximum determinant of an $n \times n$ $\{\pm 1\}$-matrix, and

$$R(n) := \frac{D(n)}{n^{n/2}} \le 1$$

be the ratio of $D(n)$ to the Hadamard bound.

Clearly $R(n) = 1$ iff n is a Hadamard order.

In this talk we consider lower bounds on $R(n)$;
these are of interest when n is not a Hadamard order.

Apart from the small cases $n \in \{1, 2\}$, Hadamard orders are multiples of four.

The *Hadamard conjecture* (actually made by Paley, 1933) is that all positive multiples of four are Hadamard orders.
This has been verified for $n < 668$.

How to find lower bounds?

There are two ways that we can obtain a lower bound on $R(n)$ if Hadamard matrices of order “close” to n exist.

- minors: Choose a Hadamard matrix H of order $h \geq n$, and take an $n \times n$ submatrix with a large determinant Δ. Theorems about minors of Hadamard matrices imply a lower bound on Δ, e.g. $h = n + 1 \Rightarrow \Delta = h^{h/2-1}$.
- bordering: Choose a Hadamard matrix H of order $h \leq n$, and add a suitable border of $d := n - h$ rows and columns. For example, if $n = 17$, we can construct a maximal determinant matrix of order 17 by choosing a Hadamard matrix of order 16 and an appropriate border.

The probabilistic method is applicable to bordering, as we can choose a border that is randomised in some way (details later).

$R(n)$ for small n

n	R	n	R	n	R	n	R
–	–	1	1	2	1	3	0.77
4	1	5	0.86	6	0.74	7	0.63
8	1	9	0.73	10	0.74	11	0.61
12	1	13	0.86	14	0.74	15	0.63
16	1	17	0.75	18	0.74	19	0.64
20	1	21	0.78	22	0.70?	23	0.61?
24	1	25	0.86	26	0.74	27	0.63?
28	1	29	0.74?	30	0.74	31	0.62?

Table: $R(n)$ for $n \leq 31$

Each block of two columns corresponds to a congruence class of n mod 4. Data from Will Orrick's website `http://www.indiana.edu/~maxdet/`.

Some known and conjectured lower bounds

Rokicki *et al* (2010) verified numerically that

$$R(n) > 1/2 \text{ for all } n \le 120,$$

and conjectured that this lower bound always holds.

However, the known rigorous bounds are much weaker than Rokicki's conjecture.

Until recently, the best published result,[1] even assuming the Hadamard conjecture, was

$$R(n) \ge \frac{1}{\sqrt{3n}}\,.$$

This bound tends to zero as $n \to \infty$.

[1] Brent and Osborn, EJC 2013.

Improved lower bounds on $R(n)$

Using the probabilistic method, we[2] recently showed that

$$R(n) \geq c_d$$

for some $c_d > 0$ that depends only on d.
Here, as usual, $d = n - h \geq 0$ is the width of the border.

For all $n \geq 1$ we have

$$R(n) \geq \left(\frac{2}{\pi e}\right)^{d/2} \left(1 - d^2 \left(\frac{\pi}{2h}\right)^{1/2}\right).$$

Also, if the Hadamard conjecture is true, then $d \leq 3$ and

$$R(n) \geq \left(\frac{2}{\pi e}\right)^{d/2} \geq \left(\frac{2}{\pi e}\right)^{3/2} > \frac{1}{9}.$$

[2] Brent, Osborn and Smith, arXiv:1402.6817, 2014.

A naive approach

How can we use the probabilistic method to bound $R(n)$?

An obvious approach is to consider a random $\{\pm 1\}$-matrix of order n, hoping that a random matrix often has a determinant close to the Hadamard bound.

In 1940, Turán showed that the

$$\mathbb{E}[\det(A)^2] = n!$$

for $\{\pm 1\}$-matrices A of order n, chosen uniformly at random.

Compare this to the Hadamard bound $\det(A)^2 \le n^n$.

$$\mathbb{E}[\det(A)^2] = n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \ll n^n.$$

This is weaker than what we need by a factor of almost e^n. Thus, the naive approach does not work.

Large determinant matrices are rare

Since we need Chebyshev's inequality later, we illustrate its use by showing that the distribution of $|\det(A)|$ has a long tail.

Theorem [Chebyshev, 1867]. Let X be a random variable with finite mean $\mu = \mathbb{E}[X]$ and finite variance $\sigma^2 = \mathbb{V}[X]$. Then, for all $\lambda > 0$,

$$\mathbb{P}[|X - \mu| \geq \lambda] \leq \frac{\sigma^2}{\lambda^2}.$$

Let $X = \det(A)$, where A is a random $\{\pm 1\}$-matrix of order n. Then $\mu = 0$ and $\sigma^2 = n!$ (by Turán's theorem). Take $\lambda = n^{n/2}/2$ (half the Hadamard bound). Then

$$\mathbb{P}\left[|\det(A)| \geq \frac{n^{n/2}}{2}\right] \leq \frac{4n!}{n^n} \sim \frac{4\sqrt{2\pi n}}{e^n}$$

is tiny if n is large.

A better approach – bordering a Hadamard matrix

Suppose $n = h + d$ where h is the order of a Hadamard matrix H, and $d \geq 0$ is small. We always choose h as large as possible, i.e. d as small as possible. If the Hadamard conjecture is true, we can assume that $0 \leq d \leq 3$.

We can start with H and add a border of d rows and columns. Since H has a large determinant (as large as possible for a $\{\pm 1\}$-matrix of order h), we hope that the resulting matrix of order n also has a large determinant.

To analyse the effect of a border on the determinant, we need to consider the *Schur complement*.

The Schur complement

Let

$$A = \begin{bmatrix} H & B \\ C & D \end{bmatrix}$$

be an $n \times n$ matrix written in block form, where H is $h \times h$ (not necessarily Hadamard, but assumed to be nonsingular), and $n = h + d > h$.

The *Schur complement* of H in A is the $d \times d$ matrix

$$D - CH^{-1}B.$$

This is relevant to our problem because it can be shown, using block Gaussian elimination, that

$$\det(A) = \det(H)\det(D - CH^{-1}B).$$

Application of the Schur complement

Let H be an $h \times h$ Hadamard matrix that is a principal submatrix of an $n \times n$ matrix A, as on the previous slide.

$$A = \begin{bmatrix} H & B \\ C & D \end{bmatrix} .$$

- Since H is Hadamard, $HH^T = hI$ and $\det(H) = h^{h/2}$, so

$$\det(A) = h^{h/2} \det(D - h^{-1} C H^T B) .$$

- In order to maximise $|\det(A)|$, we need to maximise

$$|\det(D - h^{-1} C H^T B)| .$$

A small numerical example

Suppose we want to construct a large-determinant $\{\pm 1\}$-matrix of order 5. We could start with the order 4 Hadamard matrix

$$H = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{bmatrix}$$

which has $\det(H) = 16$, and add a border along the right and bottom.

Choosing B, C, D randomly

Suppose we randomly choose B, C and D to give

$$A = \left[\begin{array}{cccc|c} +1 & +1 & +1 & +1 & -1 \\ +1 & -1 & +1 & -1 & +1 \\ +1 & +1 & -1 & -1 & +1 \\ +1 & -1 & -1 & +1 & +1 \\ \hline +1 & +1 & +1 & -1 & +1 \end{array}\right].$$

Then

$$B^T H = (H^T B)^T = [+2, -2, -2, -2],$$
$$C = [+1, +1, +1, -1],$$
$$CH^T B = 2 - 2 - 2 + 2 = 0,$$
$$\det(D - h^{-1} CH^T B) = \det(1) = 1,$$
$$\det(A) = \det(H) \cdot 1 = 16.$$

This is disappointing as $\det(A)$ is no larger than $\det(H)$.

Choosing only B randomly

Let's choose B randomly, but then choose C to avoid any cancellation in the inner product $C \cdot H^T B$, and finally choose D to maximise the resulting determinant. This gives

$$A = \left[\begin{array}{cccc|c} +1 & +1 & +1 & +1 & -1 \\ +1 & -1 & +1 & -1 & +1 \\ +1 & +1 & -1 & -1 & +1 \\ +1 & -1 & -1 & +1 & +1 \\ \hline +1 & -1 & -1 & -1 & -1 \end{array}\right]$$

In fact

$$B^T H = (H^T B)^T = [+2, -2, -2, -2],$$

$$C = [+1, -1, -1, -1],$$

$$C H^T B = 2 + 2 + 2 + 2 = 8.$$

$$\det(D - h^{-1} C H^T B) = \det(-1 - 2) = -3,$$

and $\det(A) = \det(H) \cdot (-3) = -48$. By reversing the sign of one row in A, we get the maximum possible determinant (48).

Generalisation: constructing a border for $d \geq 1$

Choose the $h \times d$ $\{\pm 1\}$-matrix B uniformly at random.

We want to choose C and D (depending on B) to maximise the expected value

$$\mathbb{E}[|\det(D - h^{-1}CH^TB)|].$$

Guided by our numerical examples, approximate this by choosing $C = (c_{ij})$, where

$$c_{ij} = \mathrm{sgn}(H^TB)_{ji} \;\text{ for }\; 1 \leq i \leq d,\; 1 \leq j \leq h$$

so there is no cancellation in the inner products defining the diagonal elements of $C \cdot H^TB$.

Finally, choose $D = -I$ (we later modify the off-diagonal elements of D get a $\{\pm 1\}$-matrix).

In the case $d = 1$ this construction is due to Brown and Spencer (1971); also (independently) to Best (1977).

Entries in the Schur complement

Write $F = h^{-1}CH^TB$, so the Schur complement is $D - F$.

The choice of D is unimportant when h is large, so for the moment we'll ignore D and concentrate on F.

- Diagonal elements. By a counting argument [Brown and Spencer 1971, Best 1977], if $h \geq 2$ then

$$\mathbb{E}[f_{ii}] = 2^{-h} \sum_{k=0}^{h} |h-2k| \binom{h}{k} = \frac{h}{2^h} \binom{h}{h/2} = \left(\frac{2h}{\pi}\right)^{1/2} + O(h^{-1/2}).$$

- Off-diagonal elements. If $i \neq j$, then

$$\mathbb{E}[f_{ij}] = 0 \text{ and } \mathbb{V}[f_{ij}] = \mathbb{E}[f_{ij}^2] = 1.$$

We expect the diagonal elements to be "large" (of order $h^{1/2}$) and the off-diagonal elements to be "small" (of order 1).

Another numerical experiment

Let's try our construction with $n = 6$, $h = 4$, $d = 2$. We choose a Hadamard matrix H of order $h = 4$ and add a border of width $d = 2$. Repeat 10^4 times, computing $F = h^{-1}CH^TB$ and $\det(F)$ each time.

In a typical experiment we find

$$\text{mean}(F) = \begin{bmatrix} 1.5002 & -0.0076 \\ -0.0002 & 1.4993 \end{bmatrix} \approx \mathbb{E}[F] = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix},$$

but

$$\text{mean}(\det(F)) = 1.6877 \neq \det(\mathbb{E}[F]) = 2.25.$$

Why the discrepancy?

The reason is that elements of F are correlated.

In particular, $\mathbb{E}(f_{12}f_{21}) \neq \mathbb{E}(f_{12})\mathbb{E}(f_{21}) = 0$.

Correlations between elements of F

From the definition of F, we see that f_{ij} depends only on the choice of columns i and j of the random border B.

Thus, f_{ij} and $f_{k\ell}$ are independent iff

$$\{i, j\} \cap \{k, \ell\} = \emptyset.$$

In the numerical example on the previous slide, f_{12} and f_{21} are correlated in a way which tends to reduce the determinant!

However, the diagonal elements f_{11} and f_{22} are independent. Thus

$$\mathbb{E}[f_{11} f_{22}] = \mathbb{E}[f_{11}]\mathbb{E}[f_{22}].$$

Inequalities for the f_{ij}

Best (1977) showed, using the Cauchy-Schwarz inequality, that

$$|f_{ij}| \leq h^{1/2}.$$

The Cauchy-Schwarz inequality also shows that, if $i \neq j$ and $k \neq \ell$, then

$$E[|f_{ij} f_{k\ell}|] \leq \sqrt{\mathbb{E}[f_{ij}^2]\mathbb{E}[f_{k\ell}^2]} = 1.$$

Using these two inequalities and the fact that the diagonal elements of F are independent, we can get a lower bound on $\mathbb{E}[\det(F)]$ by expanding the determinant as a sum of products and bounding each of the $d!$ terms.

The result is only useful if $d! \ll h$. In practice it is useful for $d \leq 3$. This is fine if you believe the Hadamard conjecture.

Example: the case $d = 2$

If $d = 2$, then

$$\det(F) = \det \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} = f_{11}f_{22} - f_{21}f_{12}.$$

Thus

$$\begin{aligned} \mathbb{E}[\det(F)] &= \mathbb{E}[f_{11}f_{22}] - \mathbb{E}[f_{21}f_{12}] \\ &\geq \mathbb{E}[f_{11}]\mathbb{E}[f_{22}] - \mathbb{E}[|f_{21}f_{12}|] \\ &\geq \frac{2h}{\pi} - O(1). \end{aligned}$$

Using the Schur complement lemma, we can deduce that $R(n) \geq \frac{2}{\pi e} > 0.23$ whenever $n - 2$ is a Hadamard order.

Previous lower bounds[3] are $\sim 5.43/n$ and $\sim 0.587/n^{1/2}$; both tend to zero as $n \to \infty$. Thus, the new bound is much better for large n.

[3] Koukouvinos, Mitrouli & Seberry [2000], Brent and Osborn [2013].

Obtaining a lower bound in the general case

To get a useful lower bound on $R(n)$ for general n, without assuming the Hadamard conjecture, we need some new ingredients:

- An upper bound on the variance of the diagonal elements of F in the probabilistic construction described above (so that we can apply Chebyshev's inequality).
- Ostrowski's inequality for determinants of matrices that are "close" to the identity matrix.
- Livinskyi's bound on gaps between Hadamard orders (to show that the result is nontrivial for all sufficiently large n).

We'll consider these in reverse order.

Gaps between Hadamard orders

From a result of Livinskyi (2012), the "gaps" between Hadamard orders near n are at most of order $n^{1/6}$, so we can assume that $d = O(h^{1/6})$.

The "error term" in our bound is $O(d^2/h^{1/2})$.
By Livinskyi's result, this is $O(1/h^{1/6})$, so $\to 0$ as $h \to \infty$.

Earlier results on gaps between Hadamard orders, by Seberry (1976) and Craigen (1995), are not sharp enough to show this.

Ostrowski's inequality

Theorem (Ostrowski, 1938). If $X = I - E$ is a $d \times d$ real matrix and the elements of E satisfy $|e_{ij}| \le \varepsilon \le 1/d$, then

$$\det(X) \ge 1 - d\varepsilon.$$

If our matrix F is close to a diagonal matrix, we can scale it to make it close to the identity matrix, and then use Ostrowski's inequality to get a lower bound on $\det(F)$.

We expect F to be close to a diagonal matrix with high probability, because the diagonal elements of F have a distribution with mean of order $h^{1/2}$ and small variance (we'll see this later), and the off-diagonal elements have mean zero and variance 1. Similarly for $G = I + F$.

The choice of D

Recall that

$$A = \begin{bmatrix} H & B \\ C & D \end{bmatrix}$$

and the Schur complement of H in A is $D - CH^{-1}B = D - F$. We choose $D = -I$ and write $G = I + F$, so $-G$ is the Schur complement.

Our choice of D is not a $\{\pm 1\}$-matrix because there are zeros off the main diagonal. However, we can later change these zeros to either $+1$ or -1 without decreasing $|\det(D - F)|$. Thus, any lower bounds on $R(n)$ that we prove using $D = -I$ are also valid for $\{\pm 1\}$-matrices.

Good G

Define a *good* G to be one for which all the g_{ij} are sufficiently close to their expected values. More precisely, g_{ij} is "good" if

$$|g_{ij} - \mathbb{E}[g_{ij}]| < d,$$

and G is "good" if all the g_{ij} are good.

The motivation for this definition is that, if G is good, we'll be able to apply Ostrowski's inequality to $\mu^{-1}G$, which is close to the identity matrix. Here $\mu = \mathbb{E}[g_{ii}] = \mathbb{E}[f_{ii}] + 1 \sim (2h/\pi)^{1/2}$.

Recall Chebyshev's inequality: $\mathbb{P}[|X - \mathbb{E}[X]| \geq \lambda] \leq \sigma^2/\lambda^2$.

This gives us a bound on the probability that an element g_{ij} is bad (i.e. not good). We take $X = g_{ij}$, $\sigma^2 = \mathbb{V}[g_{ij}]$, and $\lambda = d$. Then

$$\mathbb{P}[g_{ij} \text{ is bad}] \leq \sigma^2/d^2.$$

The off-diagonal elements

Consider the off-diagonal elements g_{ij}, $i \neq j$. For these, $\sigma^2 = 1$, so Chebyshev's inequality gives

$$\mathbb{P}[g_{ij} \text{ is bad}] \leq 1/d^2.$$

There are $d(d-1)$ off-diagonal elements, so the probability that any of them is bad is at most

$$\frac{d(d-1)}{d^2} = 1 - \frac{1}{d}\,.$$

This argument does not assume independence!

The diagonal elements

We need $V[g_{ii}]$ for a diagonal element g_{ii} of G. By a combinatorial argument, we can show that, for $h \geq 4$,

$$V[g_{ii}] = 1 + \frac{h(h-1)}{2^{h+1}} \binom{h/2}{h/4}^2 - \frac{h^2}{2^{2h}} \binom{h}{h/2}^2 .$$

Using the asymptotic expansion of $\log \Gamma(z)$ with an error bound to estimate the binomial coefficients, it follows that

$$\sigma^2 := V[g_{ii}] < 1.$$

(Is there an easier proof that avoids asymptotics?)

Chebyshev's inequality gives

$$\mathbb{P}[g_{ii} \text{ is bad}] \leq \frac{\sigma^2}{d^2} < \frac{1}{d^2} .$$

Thus, the probability that *any* diagonal element is bad is $< 1/d$.

Good G exist!

Putting the pieces together,

$$\mathbb{P}[G \text{ is bad}] < \left(1 - \frac{1}{d}\right) + \frac{1}{d} = 1.$$

Thus,

$$\mathbb{P}[G \text{ is good}] = 1 - \mathbb{P}[G \text{ is bad}] > 0.$$

Since there is a positive probability that a random choice of B gives a good G, the set of good G is nonempty!

Completing the proof

We can apply Ostrowski's inequality to $X = \mu^{-1} G$ if G is good and $\varepsilon = d/\mu$ is sufficiently small. Here $\mu = \mathbb{E}[g_{ii}] \sim \sqrt{2h/\pi}$.

More precisely, the condition on ε is $d\varepsilon < 1$, which is equivalent to $d^2 < \mu$.

This leads to the following theorem.

Theorem. If $n = h + d$ where $d \geq 1$ and there exists a Hadamard matrix of order $h \geq 4$, then

$$D(n) \geq h^{h/2} \mu^d (1 - d^2/\mu).$$

Note. Since μ is of order $h^{1/2} \approx n^{1/2}$ and $d \ll n^{1/6}$ [Livinskyi],

$$d^2/\mu \ll n^{1/3}/n^{1/2} = 1/n^{1/6} \to 0,$$

so $d^2 < \mu$ for all sufficiently large n.

The lower bound on $D(n)$

In the inequality

$$D(n) \geq h^{h/2} \mu^d (1 - d^2/\mu)$$

- the factor $h^{h/2}$ comes from the determinant of H,
- the factor μ^d comes from the expected product of the diagonal elements of G, and
- the factor $(1 - d^2/\mu)$ comes from the application of Ostrowski's inequality.

The first two factors seem unavoidable. The last factor can be improved by using the "Lovász Local Lemma" and Hoeffding's tail inequality instead of Chebyshev's inequality.[4]

[4] Brent, Osborn and Smith, arXiv:1402.6817.

A lower bound on $R(n)$

Corollary. If $d \geq 1$, $n = h + d$ as above, then

$$R(n) \geq \left(\frac{2}{\pi e}\right)^{d/2} \left(1 - d^2\sqrt{\frac{\pi}{2h}}\right).$$

Since $d^2/h^{1/2} = O(1/n^{1/6})$, this is close to the bound

$$R(n) \geq \left(\frac{2}{\pi e}\right)^{d/2}$$

that we can prove for $d \leq 3$.

We[5] can also prove the latter inequality if $n \geq n_0$, where n_0 is an absolute constant (independent of d).

A plausible conjecture is that the inequality holds for all positive n.

[5] BOS, arXiv:1402.6817.

References

N. Alon and J. H. Spencer, *The Probabilistic Method*, 3rd edn., Wiley, 2008.

M. R. Best, The excess of a Hadamard matrix, *Indag. Math.* 39 (1977), 357–361.

R. P. Brent and J. H. Osborn, General lower bounds on maximal determinants of binary matrices, *Electron. J. Comb.* 20, 2 (2013), #P15.

R. P. Brent, J. H. Osborn and W. D. Smith, Probabilistic lower bounds on maximal determinants of binary matrices, arXiv:1501.06235v4, 26 Jan. 2016.

T. A. Brown and J. H. Spencer, Minimization of ±1 matrices under line shifts, *Colloq. Math.* 23 (1971), 165–171 & 177.

P. Erdős and J. Spencer, *Probabilistic Methods in Combinatorics*, Akadémiai Kiadó, Budapest, 1974.

J. Hadamard, Résolution d'une question relative aux déterminants, *Bull. des Sci. Math.* 17 (1893), 240–246.

C. Koukouvinos, M. Mitrouli and J. Seberry, Bounds on the maximum determinant for $(1, -1)$ matrices, *Bull. Inst. Combinatorics and its Applications* 29 (2000), 39–48.

I. Livinskyi, *Asymptotic existence of Hadamard matrices*, M.Sc. thesis, University of Manitoba, 2012. http://hdl.handle.net/1993/8915

A. M. Ostrowski, Sur l'approximation du déterminant de Fredholm par les déterminants des systèmes d'equations linéaires, *Ark. Math. Stockholm* 26A (1938), 1–15.

T. Rokicki, I. Kazmenko, J-C. Meyrignac, W. P. Orrick, V. Trofimov and J. Wroblewski, *Large determinant binary matrices: results from Lars Backstrom's programming contest*, unpublished report, July 31, 2010. See also http://www.recmath.org/contest/Determinant/matrix.html