Lower bounds on maximal determinants via the probabilistic method

Richard P. Brent
ANU

14 December 2012

joint work with

Warren Smith and Judy-anne Osborn

Presented at the 2012 ACCMCC conference, UNSW

Introduction – the Hadamard bound and conjecture

- $D(n) :=$ denote the maximum determinant attainable by an $n \times n$ $\{\pm 1\}$-matrix.
- Hadamard proved the upper bound $D(n) \leq n^{n/2}$.
- A *Hadamard matrix* of order n is an $n \times n$ $\{\pm 1\}$-matrix A with $\det(A) = \pm n^{n/2}$.
- If a Hadamard matrix of order n exists, then $n = 1$, 2, or a multiple of 4.
- The *Hadamard conjecture* is that Hadamard matrices exist for every positive multiple of 4.
- This talk is about lower bounds on $D(n)$.

Notation

- $\mathcal{H}$ is the set of all possible orders of Hadamard matrices.
- $\mathcal{R}(n) := D(n)/n^{n/2}$.
  The Hadamard bound is $\mathcal{R}(n) \le 1$.
  We are interested in lower bounds on $\mathcal{R}(n)$.
- $d := n - \max\{h \in \mathcal{H} \mid h \le n\}$.
  In other words, $n = h + d$, $d \ge 0$, and $h \in \mathcal{H}$ is maximal.
- To avoid trivial cases, assume that $n \ge h \ge 4$.
- We'll use Vinogradov's notation:
  $f \ll g$ means $f = O(g)$ and $f \gg g$ means $g = O(f)$.
- $f = O_d(g)$ or $f \ll_d g$ means that the implied "constant" depends on d (so it is only constant if d is fixed).

Previous results

For those of you who attended my AustMS talk in Ballarat – the problem is the same, but the results are generally better!

In all previous papers that we are aware of (including our own), general lower bounds on $\mathcal{R}(n)$ tend to zero as $n \to \infty$, unless $n \in \mathcal{H}$ or $n - 1 \in \mathcal{H}$.

For example, de Launey and Levin (2009) showed that

$$\mathcal{R}(n) \geq \frac{2^{1/2}e}{n}\left(1 + O\left(\frac{1}{n}\right)\right)$$

if $n \equiv 2 \pmod 4$, assuming the Hadamard conjecture.

Under the same assumption, our new result is

$$\mathcal{R}(n) > \frac{2}{\pi e} \approx 0.2342$$

Previous approaches

The most successful previous approaches to obtaining general lower bounds (as opposed to bounds for specific small values of n) used either bordering or minors.

- bordering: choose a Hadamard matrix H of order $h < n$, and add a border of $n - h$ rows and columns to H.
- minors: choose a Hadamard matrix H of order $h > n$, and consider some $n \times n$ submatrix of H.

The best lower bound obtained via bordering or minors was

$$\mathcal{R}(n) \gg n^{-\delta/2} \text{ where } \delta = |n - h|$$

[Koukovinos, Mitrouli and Seberry; de Launey and Levin; Brent and Osborn] with one exception (next slide).

Improved bound for bordering if $n = h + 1$

For $n - h = 1$, the lower bound can be improved to

$$\mathcal{R}(n) \geq \text{ constant}$$

by using a probabilistic method due to Brown and Spencer (1971), Erdős and Spencer (1974), and Best (1977).

The idea is to add a border of one row and column to a Hadamard matrix in a (semi-)probabilistic manner that gives a large determinant (on average).

Curiously, none of these authors seems to have considered adding a larger border.

Our approach

We generalise the probabilistic bordering method by taking a Hadamard matrix of order $h < n$ and adding a border of $d = n - h$ rows and columns in a (semi-) probabilistic manner.

This enables us to obtain lower bounds of the form

$$\mathcal{R}(n) \geq \kappa_d > 0,$$

where κ_d depends only on d.

For example,

$$\mathcal{R}(n) \geq 0.07\,(0.352)^d > 3^{-(d+3)}.$$

The Schur complement

Let

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

be an $n \times n$ matrix written in block form, where A is $h \times h$, and $n = h + d > h$.

The *Schur complement* of A in $\widetilde{A}$ is the $d \times d$ matrix

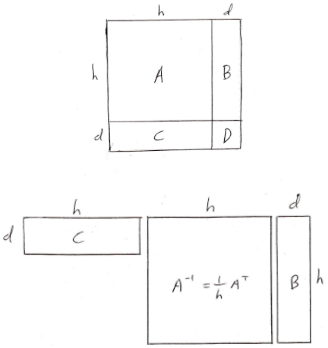
$$D - CA^{-1}B.$$

The Schur complement is relevant to our problem because

$$\det(\widetilde{A}) = \det(A)\det(D - CA^{-1}B).$$

The Schur complement is not in general a $\{\pm 1\}$-matrix.

The block matrix $\widetilde{A}$ and Schur complement



$$\det(\widetilde{A}) = \det(A)\det(D - CA^{-1}B).$$

Application of the Schur complement

Take A to be an $h \times h$ Hadamard matrix that is a principal submatrix of an $n \times n$ matrix, $n = h + d$.

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

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

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

- The problem is to maximise the order d determinant

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

Using the probabilistic method

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

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

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

Approximate this by choosing $C = (c_{ij})$, where

$$c_{ij} = \mathrm{sgn}(A^T B)_{ji} \text{ for } 1 \le i \le d,\ 1 \le j \le h$$

so there is no cancellation in the inner products defining the diagonal elements of $C \cdot A^T B$.

For $d = 1$ this is the same as the choice made by Best, Brown, Erdős and Spencer.

Entries in the Schur complement

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

The choice of D is not important (at least as $h \to \infty$), so for simplicity we'll ignore D and concentrate on F.

- Diagonal elements. By a counting argument [Best *et al*]

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

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

$$E(f_{ij}) = 0 \text{ and } E(f_{ij}^2) = 1.$$

- All elements. $|f_{ij}| \le h^{1/2}$.

The contribution of the off-diagonal elements

We want to approximate the determinant of the Schur complement by the product of its diagonal elements.

One way of showing that the contribution from the off-diagonal elements is (usually) small is to use the Cauchy-Schwarz inequality:

$$E(|f_{ij}f_{k\ell}|) \le \sqrt{E(f_{ij}^2)E(f_{k\ell}^2)} = 1.$$

We can not assume that f_{ij} and $f_{k\ell}$ are independent, even if $i \neq j$ and $k \neq \ell$. For example, f_{12} and f_{21} are dependent.

Exercise. Show that f_{ij} depends only on columns i and j of B. Deduce that f_{ij} and $f_{k\ell}$ are independent iff $\{i, j\} \cap \{k, \ell\} = \emptyset$.

Using Cauchy-Schwartz to estimate $\det(F)$

We want a lower bound on $E(\det(F))$ for fixed d and large h.

For example, if $d = 3$,

$$\det(F) = \det \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} = f_{11}f_{22}f_{33} + \text{other terms},$$

and a typical "other term" has expectation $O(h^{1/2})$ as

$$|E(f_{12}f_{21}f_{33})| \le E(|f_{12}f_{21}|)\max(|f_{33}|) \le h^{1/2}.$$

Thus, using independence of f_{11}, f_{22} and f_{33},

$$E(\det(F)) = E(f_{11}f_{22}f_{33}) + O_d(h^{1/2}) = \left(\frac{2h}{\pi}\right)^{3/2} + O_d(h^{1/2}).$$

First result

Theorem. If $d \geq 1$, $h \in \mathcal{H}$, $n = h + d$, and $h \geq h_0(d)$, then

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

The constant $2/(\pi e)$ appearing here is nice, but probably not best possible, since our proof uses expectations, not maxima.

From the Barba and Ehlich-Wojtas upper bounds, we know that

$$\limsup_{\mathcal{H} \ni h \to \infty} \mathcal{R}(h+d) \leq \left(\frac{2}{e}\right)^{d/2} \quad \text{for} \quad d \leq 2.$$

Small d

For $0 \le d \le 3$, our theorem implies, after considering the cases with $h < h_0(3)$ separately, that

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

Numerically,

$$\left(\frac{2}{\pi e}\right)^{1/2} > 0.4839 \;\text{ so }\; \mathcal{R}(n) \ge (0.4839)^d.$$

If the Hadamard conjecture is true, then every positive integer divisible by 4 is a Hadamard order, so $0 \le d \le 3$, and the inequality always holds.

Ameliorating the cutoff $h_0(d)$

If the Hadamard conjecture is false, we have to consider $d \geq 4$.

Our theorem required $h \geq h_0(d)$, where $h_0(d)$ grows too fast for comfort, roughly as

$$(d/2)^{2d}.$$

We can reduce (and even eliminate) the cutoff $h_0(d)$ by using a different way to bound the effect of off-diagonal elements in the Schur complement.

The idea is to use a Chernoff/Hoeffding tail inequality, combined with a lower bound on the determinant of a diagonally dominant matrix.

There is a price to pay – the proof is more complicated, and the final inequality that we get is slightly weaker.

Using Hoeffding's tail inequality

- Let $X_1, \ldots, X_h$ be independent random variables with sum Y, where $X_i \in [a_i, b_i]$. Then, for all $t > 0$,

$$\Pr\left(|Y - E[Y]| \geq t\right) \;\leq\; 2\,\exp\left(\frac{-2t^2}{\sum_{i=1}^{h}(b_i - a_i)^2}\right) .$$

- This can be applied with $Y = f_{ij}$, which can be written as a sum of h bounded, independent random variables.
- If the off-diagonal elements of the Schur complement are usually small and the diagonal elements are often large, then with positive probability we can use a lower bound on the determinant of a diagonally dominant matrix.

Second result

We can remove the restriction on h at the cost of reducing the constant from $\left(\frac{2}{\pi e}\right)^{1/2} \approx 0.484$ to $0.352 > 1/3$.

Theorem. If $d \geq 0$, $h \in \mathcal{H}$, and $n = h + d$, then

$$\mathcal{R}(n) > 3^{-(d+3)}.$$

Comparison: the bound of Clements and Lindström (1965) is

$$\mathcal{R}(n) > (3/4)^{n/2}.$$

Our bound is much sharper since $d \ll n^{1/6}$ [Livinskyi 2012]. It is sharper than the bounds of Koukouvinos, Mitrouli and Seberry (also de Launey and Levin, Brent and Osborn) if $d > 0$ is fixed and $n \to \infty$; all these bounds are at best $\mathcal{R}(n) \gg n^{-1/2}$.

Ingredients in the proof

The proof uses

- Hoeffding's tail inequality for a sum of bounded independent random variables,
- a new (best possible) lower bound on the determinant of a diagonally dominant matrix, improving on what can be obtained from Gerschgorin's theorem,
- various known constructions for Hadamard matrices,
- results of Livinskyi (2012) on the asymptotic density of Hadamard matrices, and
- a computer-aided analysis of a set of 32 exceptional cases with $n < 60480$.

For the details, see arXiv:1211.3248.

Conjecture

We conjecture that

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

Evidence. The conjecture holds for:

- for $0 \leq d \leq 3$ (implied by the Hadamard conjecture),
- for all $d \geq 0$ if $n \geq n_0(d)$ is sufficiently large,
- for all $n \leq 120$ (in fact $\mathcal{R}(n) > 1/2$ for $n \leq 120$),
- for many larger values of n for which we have computed a lower bound on $\mathcal{R}(n)$ using a probabilistic algorithm based on our construction.

Acknowledgements

Thanks to

- Dragomir Djoković and Ilias Kotsireas for sharing their list of known Hadamard orders;
- Robert Craigen for informing us of the work of his student Ivan Livinskyi; and
- Will Orrick for his comments.

References

N. Alon and J. H. Spencer, *The Probabilistic Method,* third edition, 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, submitted. Also arXiv:1208.1805v3, 4 Sept. 2012.

R. P. Brent, J. H. Osborn and W. D. Smith, Lower bounds on maximal determinants of ± 1 matrices via the probabilistic method. arXiv:1211.3248v2, 3 Dec. 2012.

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

References continued

G. F. Clements and B. Lindström, A sequence of (± 1)-determinants with large values, *Proc. Amer. Math. Soc.* **16** (1965), 548–550.

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

S. Gerschgorin, Über die Abgrenzung der Eigenwerte einer Matrix, *Izv. Akad. Nauk. USSR* **6** (1931), 749–754.

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

W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Ass.* **58** (1963), 13–30.

References continued

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

W. de Launey and D. A. Levin, $(1, -1)$-matrices with near-extremal properties, *SIDMA* **23** (2009), 1422–1440.

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

R. S. Varga, *Geršgorin and His Circles*, Springer Series in Computational Mathematics, Vol. 36, 2004.

J. Seberry Wallis, On the existence of Hadamard matrices, *J. Comb. Theory* **21** (1976), 188–195.