Improved lower bounds on the Hadamard maxdet problem, Part II

Richard P. Brent
ANU and Newcastle

1 October 2013

joint work with

Judy-anne Osborn
Warren D. Smith

Presented at the 57th annual meeting of the Australian Mathematical Society, Sydney

Outline – Part II

- Recap of Part I
- A small example
- Determinant of a perturbation of the identity
- Inequalities of Chebyshev and Cantelli
- A probabilistic lower bound using Chebyshev's inequality
- The Lovász Local Lemma
- Hoeffding's inequality
- Sharper lower bounds
- Limitations of the probabilistic approach
- Numerical example
- A conjecture

Recalling Part I

This is Part II of a combined talk. Here is a quick summary of Part I.

H is a Hadamard matrix of order h. We (probabilistically) add a border of d rows and d columns so that the $n \times n$ $\{\pm 1\}$-matrix

$$\begin{pmatrix} H & B \\ C & D \end{pmatrix}$$

has large (expected) determinant. This gives us a lower bound on the *maximal determinant* function $D(n)$. $(n = h + d)$

Since $|\det(H)| = h^{h/2}$ is fixed, this amounts to choosing the border (B, C and D) so that the Schur complement $D - CH^{-1}B$ has a large determinant. Note that $H^{-1} = h^{-1}H^T$.

We define $F := CH^{-1}B = h^{-1}CH^TB$ and $G := F + I$.

Determinant of the Schur complement

We are interested in the determinant Δ of the Schur complement $D - h^{-1}CH^TB = D - F$, where H, B, C and D are $\{\pm 1\}$-matrices.

We can always choose D so that

$$|\det(D - F)| \geq |\det(F + I)| = |\det(G)|\,.$$

Thus, there is no harm in assuming that $D = -I$ since this will give valid lower bounds on $|\Delta|$ (even though $-I$ is not a $\{\pm 1\}$-matrix). In the following we consider $G = F + I$.

The diagonal elements g_{ii} of G are expected to be of order $h^{1/2}$, and the off-diagonal elements of order unity, so $h^{-1/2}G$ is expected to be a perturbation of the $d \times d$ identity matrix.

Dependencies in the Schur complement

With our probabilistic construction, the elements of the matrix $F = h^{-1}CH^TB$ are not independent. (If they were, the lower bound proofs would be much easier!)

However, from the construction, f_{ij} depends only on columns i and j of the random matrix B. Thus, f_{ij} and $f_{k\ell}$ are independent whenever $\{i, j\} \cap \{k, \ell\} = \emptyset$.

Note that the diagonal elements f_{ii} are mutually independent, as f_{ii} depends only on column i of B.

Similar remarks apply to $G = F + I$.

A small example ($h = 4$, $d = 2$, $n = 6$)

Consider the case $n = 6$. It is known that $D(6) = 160 = 10 \times D(4)$, so the Schur complement determinant Δ satisfies $|\Delta| \le 10$ (achievable).

Writing the matrix entries as $\mathbb{E}[g_{ij}] \pm \mathbb{V}[g_{ij}]^{1/2}$, the probabilistic construction gives

$$G \approx \begin{pmatrix} 2.5 \pm 0.5 & 0.0 \pm 1.0 \\ 0.0 \pm 1.0 & 2.5 \pm 0.5 \end{pmatrix}.$$

Here $\mathbb{E}[g_{11}g_{22}] = \mathbb{E}[g_{11}]\mathbb{E}[g_{22}] = 6.25$ (they are independent), but $\mathbb{E}[\det(G)] = \mathbb{E}[g_{11}g_{22} - g_{12}g_{21}] \approx 5.69 < 6.25$ as $\mathbb{E}[g_{12}g_{21}] \approx 0.56 \ne 0$ (g_{12} and g_{21} are not independent).

The off-diagonal elements of G conspire against us to reduce $\mathbb{E}[\det(G)]$ from what would be expected if we just considered the diagonal elements of G.

This motivates the following Lemma.

A determinantal inequality

Lemma

If $E \in \mathbb{R}^{d \times d}$, $|e_{ij}| \le \varepsilon$ *for* $1 \le i, j \le d$, *and* $d\varepsilon \le 1$, *then*

$$\det(I - E) \ge 1 - d\varepsilon\,.$$

Proof. See [BOS, arXiv:1211.3248v3, Lemma 8]. □

Remark. The Lemma is best possible, since it follows from a well-known rank-1 update formula that

$$\det(I - \varepsilon e e^T) = 1 - d\varepsilon\,.$$

Gerschgorin's theorem gives the weaker inequality

$$\det(I - E) \ge (1 - d\varepsilon)^d\,.$$

Inequalities of Chebyshev and Cantelli

Let X be a random variable with finite mean μ and standard deviation $\sigma = \mathbb{V}[X]^{1/2} > 0$.

Chebyshev's inequality says that, for any positive λ,

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

Cantelli's inequality is analogous but one-sided:

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

and by symmetry

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

Notation: μ and σ^2

In the following we assume $h \geq 4$.

$$\mu := \mathbb{E}[g_{ii}] = \mathbb{E}[f_{ii}] + 1$$

is the expectation of the diagonal elements of G. From Part I,

$$\mu = 1 + 2^{-h} h \binom{h}{h/2} > \left(\frac{2h}{\pi}\right)^{1/2} .$$

Also,

$$\sigma^2 := \mathbb{V}[g_{ii}]$$

is the variance of the diagonal elements. From Part I,

$$0.045 \approx 1 - 3/\pi < \sigma^2 \leq 1/4 .$$

The upper bound $1/4$ is attained at $h = 4$,
and the lower bound $1 - 3/\pi$ is the limit as $h \to \infty$.

A new lower bound for $D(n)$

Theorem

Suppose $n = h + d$ where $d \geq 0$ and $h \geq 4$ is a Hadamard order. Then

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

Remarks

By a result of Livinskyi (2012) on gaps between Hadamard orders, $d = O(h^{1/6})$. Thus

$$\left(1 - d^2\sqrt{\frac{\pi}{2h}}\right) = 1 - O(n^{-1/6}) \to 1 \text{ as } n \to \infty.$$

Lower bound for $R(n)$

Corollary

$$R(n) \geq \left(\frac{2}{\pi e}\right)^{d/2} \left(1 - O(n^{-1/6})\right) \text{ as } n \to \infty\,.$$

Remark

The factor $(1 - O(n^{-1/6}))$ can be omitted if $d \leq 3$.
We conjecture that it can always be omitted.

Idea of proof of the Theorem

The idea is to choose B uniformly at random, and say that the choice is *good* if the resulting matrix $G = I + h^{-1}CH^T B$ is "close" to the diagonal matrix μI in the sense that all the elements of $\mu^{-1}G - I$ are sufficiently small.

If the probability of a good choice is positive, then a good choice must exist, and we obtain a lower bound from the determinantal lemma (if it is applicable).

The probability of a good choice can be bounded using Chebyshev's inequality and our results on $\mathbb{E}[g_{ij}]$ and $\mathbb{V}[g_{ij}]$.

Sketch of proof

Let λ be a positive parameter to be chosen later. Using Chebyshev's inequality, for the off-diagonal elements with variance 1,

$$\mathbb{P}[|g_{ij}| \geq \lambda] \leq 1/\lambda^2 .$$

For the diagonal elements with variance $\sigma^2 \leq 1/4$,

$$\mathbb{P}[|g_{ii} - \mu| \geq \lambda] \leq \sigma^2/\lambda^2 .$$

If

$$d(d-1) \cdot \mathbb{P}[|g_{ij}| \geq \lambda] + d \cdot \mathbb{P}[|g_{ii} - \mu| \geq \lambda] < 1, \quad (*)$$

then there is a positive probability that none of the blue inequalities hold. $(*)$ holds if $\lambda = d$. With positive probability we can apply the determinantal lemma with $\varepsilon = \mu^{-1} d$ to $\mu^{-1} G$ (provided $d\varepsilon \leq 1$, i.e. $d^2 \leq \mu$, so ε is sufficiently small).

Sketch of proof (continued)

With positive probability,

$$\det(\mu^{-1}G) \geq 1 - d\varepsilon = 1 - d^2/\mu\,.$$

This is equivalent to

$$\det(G) \geq \mu^d(1 - d^2/\mu)\,.$$

The theorem follows from the Schur complement lemma, as

$$\left|\det\begin{pmatrix} H & B \\ C & D \end{pmatrix}\right| \geq |\det(H)| \cdot |\det(G)| = h^{h/2}|\det(G)|$$

for some choice of the $\{\pm 1\}$-matrix D. ∎

What if $h < \pi d^4/2$?

The Theorem is trivial if $\mu \le d^2$, as then $(1 - d^2/\mu) \le 0$ and we don't get any useful information.

Since $\mu \sim (2h/\pi)^{1/2}$, this means that the Theorem is only useful when $h \ge \pi d^4/2$ (approx.), or roughly $d = O(h^{1/4})$.

In this situation we can apply the construction with random B and see what happens. In all the cases that we have tried, a few random trials are sufficient to find a matrix G such that

$$\det(G) \ge \mu^d,$$

so we can ignore the factor $(1 - d^2/\mu)$ in the Theorem.

There are some theoretical improvements that go some way (but not all the way) towards justifying this. We'll outline them if time permits.

The Lovász Local Lemma

We need to state the *Lovász Local Lemma* [Erdős and Lovász, 1975].

Lemma (Lovász Local Lemma, symmetric case)

Let $E_1, E_2, \ldots E_m$ be events in an arbitrary probability space. Suppose that each event E_i is mutually independent of all the other events E_j except for at most D of them, and that $\mathbb{P}[E_i] \le p$ for $1 \le i \le m$. If

$$ep(D+1) \le 1$$

then $\mathbb{P}[\bigwedge_{i=1}^{m} \overline{E_i}] > 0$. (In other words, with positive probability none of the events E_i hold.)

Counting dependencies in the Schur complement

- We noted previously that f_{ij} and $f_{k\ell}$ are independent whenever $\{i,j\} \cap \{k,\ell\} = \emptyset$.
- Assume that $d > 1$. There are $4d - 4$ entries in the union of rows i and j and columns i and j of F.
- Thus, f_{ij} is dependent on at most $4d - 5$ of the other $f_{k\ell}$. We can apply the Lovász Local Lemma with $D = 4d - 5$.
- Instead of $\lambda = d$ we can take $\lambda = \sqrt{e(D+1)}$ in the proof of the theorem. This changes the $1 - d^2/\mu$ term in the lower bound to $1 - O(d^{3/2}/\mu)$. Thus, the result is nontrivial if $d = O(h^{1/3})$ instead of the previous (stricter) condition $d = O(h^{1/4})$.
- The resulting bound is sharper for $d \geq 10$.

Hoeffding's tail inequality

Hoeffding's tail inequality applies for sums of independent, bounded random variables.

Theorem (Hoeffding, 2-sided version)

Let $X_1, \ldots, X_h$ be independent random variables with sum $Y = X_1 + \cdots + X_h$. Assume that $X_i \in [a_i, b_i]$. Then, for all $t > 0$,

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

This can be applied to the off-diagonal elements f_{ij} since they may be written as sums of h independent random variables.

Note that the bound is exponentially decreasing.

Compare Chebyshev's inequality, where the bound is polynomially decreasing.

Another improvement

Using Cantelli's inequality for the diagonal elements of G, and Hoeffding's inequality for the off-diagonal elements, and allowing different tolerances for the diagonal and off-diagonal elements (which requires a generalisation of the determinantal lemma), we can replace the $d^2/\mu = O(d^2/h^{1/2})$ term by $O(d^{5/3}/h^{2/3})$.

Now the result is nontrivial for $d = O(h^{2/5})$ (compare $d = O(h^{1/3})$ using the Lovász Local Lemma).

These improvements are significant for small h, but they do not increase the main factor of order

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

in the lower bounds.

Limitations of the probabilistic approach

The Barba and Wojtas constructions show that, in the cases $d = 1$ and $d = 2$ respectively,

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

as $n \to \infty$ in a certain infinite sequence of values for which the Barba/Wojtas upper bounds are attained.

In contrast, the probabilistic method gives a lower bound

$$\sim \left(\frac{2}{\pi e}\right)^{d/2} .$$

The factor $\pi^{-d/2}$ in the lower bound seems to be an artefact of the probabilistic method – we are actually estimating the mean determinant in a certain ensemble of matrices instead of the maximum determinant.

Another limitation

In cases where we know the maximal determinant matrices of order n (that is, for $n \leq 21$ and a sparse set of larger n), it is not always true that a maximal determinant matrix contains a Hadamard matrix of order $4\lfloor n/4 \rfloor$.

Examples are $n = 13, 14, 15, 18, 19, 21$. In such cases our construction must underestimate $D(n)$.

Numerical example

Consider the case $n = 668$. It is not known if a Hadamard matrix of this order exists.

We can take $h = 664$, $d = 4$. Then $\mu \approx 21.55$, $\sigma^2 \approx 0.0464$.

Our first Theorem gives $\det(G)/\mu^d \geq 0.2576$.

For comparison, the best known deterministic construction (based on bordering) gives $\det(G)/\mu^d$ of order $1/n^2 < 10^{-5}$.

Using the Lovász Local Lemma does not help as $d < 10$.

Using Cantelli's and Hoeffding's inequalities with optimal choices of the two parameters (the diagonal and off-diagonal tolerances) gives $\det(G)/\mu^d \geq 0.7990$.

The best we can expect from the probabilistic approach is $\det(G)/\mu^d \geq 1$.

Conjecture

We conjecture that

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

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 $R(n) > 1/2$ for $n \leq 120$);
- for many larger values of n for which we have computed a lower bound on $R(n)$ using a probabilistic algorithm based on our construction.

References

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

G. Barba, Intorno al teorema di Hadamard sui determinanti a valore massimo, *Giorn. Mat. Battaglini* **71** (1933), 70–86.

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, *Electronic J. of Combinatorics* **20**(2), 2013, #P15, 12 pp.

R. P. Brent and J. H. Osborn, *Note on a double binomial sum relevant to the Hadamard maximal determinant problem*, 12 Sept. 2013, arXiv:1309.2795v2.

R. P. Brent, J. H. Osborn and W. D. Smith, *Lower bounds on maximal determinants of* ± 1 *matrices via the probabilistic method*, 5 May 2013, arXiv:1211.3248v3.

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

References cont

H. Ehlich, Determinantenabschätzungen für binäre Matrizen, *Math. Z.* **83** (1964), 123–132; *ibid* **84** (1964), 438–447.

H. Enomoto and M. Miyamoto, On maximal weights of Hadamard matrices, *J. Combin. Theory* A **29** (1980), 94–100.

P. Erdős and J. Spencer, *Probabilistic Methods in Combinatorics*, Academic Press, New York, 1974.

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

I. Livinskyi, *Asymptotic existence of Hadamard matrices*, M.Sc. thesis, University of Manitoba, 2012.

K. W. Schmidt and E. T. H. Wang, The weights of Hadamard matrices, *J. Combin. Theory* A **23** (1977), 257–263.

W. Wojtas, On Hadamard's inequality for the determinants of order non-divisible by 4, *Colloq. Math.* **12** (1964), 73–83.