Lower Bounds for the Hadamard Maximal Determinant Problem

Richard P. Brent
Australian National University

7 February 2015

Gene Golub memorial lecture presented at Hong Kong Baptist University

Joint work with
Warren Smith and Judy-anne Osborn

Abstract

Gene Golub was interested in both matrix computations and statistics. In this Golub memorial lecture I will consider a problem that involves aspects of both – the *Hadamard maximal determinant problem*.

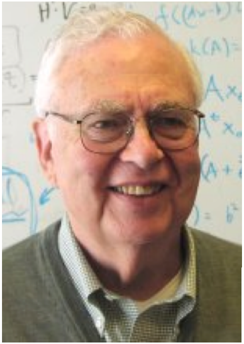
The problem is to find the maximal determinant of an $n \times n$ matrix whose elements are in $[-1, 1]$. A matrix achieving the maximum is known as a *D-optimal design* and has applications in the design of experiments. Hadamard proved an upper bound $n^{n/2}$ on the determinant, but his upper bound is not achievable for every positive integer n. For example, if $n = 3$ then Hadamard's upper bound is $3\sqrt{3} \approx 5.2$, but the best that can be achieved is 4.

Abstract cont.

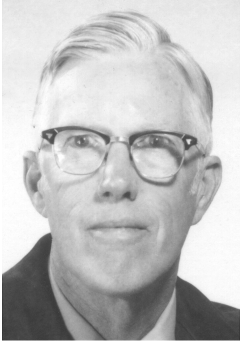
A *Hadamard matrix* is an $n \times n$ matrix that achieves Hadamard's bound. The *Hadamard conjecture* is that a Hadamard matrix exists whenever n is a multiple of four. I will consider how close to Hadamard's bound we can get when n is *not* the order of a Hadamard matrix, and outline a recent proof that Hadamard's bound is within a constant factor of the best possible, provided n is close (in a sense that will be made precise) to the order of a Hadamard matrix. In particular, if the Hadamard conjecture is true, then the constant factor is at most $(\pi e/2)^{3/2}$.

This is joint work with Judy-anne Osborn and Warren Smith.

Gene Golub and George Forsythe



Gene H. Golub (1932–2007)

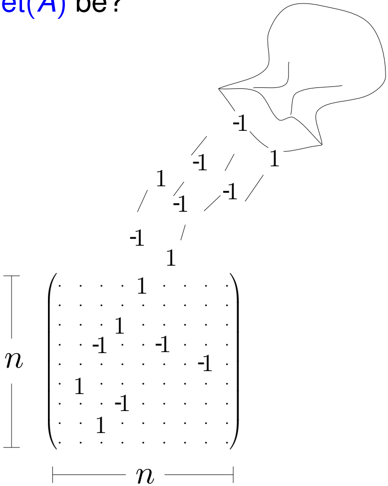


George E. Forsythe (1917–1972)

Gene Golub in 2007 at *Stanford 50* – a conference celebrating the 50th anniversary of George Forsythe's arrival at Stanford and the 75th birthday (including non-leap years) of Gene Golub. This was the last time that I saw Gene.

The Hadamard maximal determinant problem

Suppose A is an $n \times n$ matrix with entries in $\{-1, +1\}$ (we'll call this a "$\{\pm 1\}$-matrix of order n").
How large can $\det(A)$ be?



Hadamard's upper bound

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 (e.g. all powers of two). Such n are called *Hadamard orders* and the matrices attaining the bound are called *Hadamard matrices*.

Desplanques, Lévy, Muir, Sylvester, Thomson (Lord Kelvin), and others also made contributions.

Jacques Hadamard

Jacques Hadamard (1865–1963)

A short proof of Hadamard's inequality

Consider the "Gram matrix" $G = A^T A$. Note that G is positive semi-definite, so has non-negative real eigenvalues λ_j. Also, $\mathrm{diag}(G) = nI$, so $\mathrm{trace}(G) = n^2$. Thus

$$
\begin{aligned}
|\det(A)|^{2/n} = \det(G)^{1/n} &= \left(\prod_j \lambda_j\right)^{1/n} \\
&\le \frac{1}{n}\sum_j \lambda_j \quad \text{(by the AGM inequality)} \\
&= \frac{\mathrm{trace}(G)}{n} = n.
\end{aligned}
$$

Thus $|\det(A)| \le n^{n/2}$, and there is equality iff $G = nI$ (because the AGM inequality is strict unless all the λ_j are equal). $\square$

The proof shows that Hadamard matrices are orthogonal (up to a scale factor), in fact $A^T A = AA^T = nI$.

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.
- 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.
- We can ask the same question for $(n-1) \times (n-1)$ matrices whose entries are in $\{0, 1\}$. The answer is the same, except for a scaling factor of 2^{n-1} (next slide).

Determinants of $\{\pm 1\}$-matrices and $\{0,1\}$-matrices

An $n \times n$ $\{\pm 1\}$-matrix always has determinant divisible by 2^{n-1}, because of a well-known mapping from $\{0,1\}$-matrices of order $n-1$ to $\{\pm 1\}$-matrices of order n.

The mapping is reversible if we are allowed to normalise the first row and column of the $\{\pm 1\}$-matrix by changing the signs of rows/columns as necessary.

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{\text{double}} \begin{pmatrix} 2 & 0 & 2 \\ 2 & 2 & 0 \\ 0 & 2 & 2 \end{pmatrix}$$

$$\xrightarrow{\text{border}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix} \xrightarrow[\text{first row}]{\text{subtract}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \end{pmatrix}$$

Design of experiments

The field of *Design of Experiments* was pioneered by Charles Sanders Peirce (in the period 1877–1883) and later developed by Ronald Aylmer Fisher (around 1926–1935).

Suppose we want to perform m experiments to find information about the effect of n variables, where $m \geq n$. For example, we could be trying to estimate the weights of n objects using m weighings, or estimate the effect of n different drugs on m patients. We can model the experiment by an $m \times n$ matrix A of $\{0, \pm 1\}$ entries.

Provided the outcomes are linear functions of the variables, a sensible criterion to choose the best experimental design is to maximize $\det(A^T A)$. Here the Gram matrix $A^T A$ is called the *information matrix* of the design.

An $m \times n$ $\{\pm 1\}$-matrix A for which $\det(A^T A)$ is maximal is called a *D-optimal design*, and if $m = n$ it is called *saturated*.

Charles S. Peirce and Ronald A. Fisher

Charles S. Peirce (1839–1914)

Ronald A. Fisher (1890–1962)

Other criteria

Several other design criteria have been suggested. One that would be close to Gene Golub's heart is *E-optimal design*, which seeks to maximise the smallest eigenvalue of the information matrix – equivalently, maximise the smallest singular value of A.

Gene's numberplate

In this talk I will only consider D-optimal design, which maximises the product of singular values of A, or (equivalently) maximises the differential Shannon information content of the parameter estimates.

The Hadamard conjecture

It is conjectured that Hadamard matrices exist for orders $n = 1, 2$, and $4k$ for all positive integers k (it is easy to prove that these are the only possible orders). This conjecture is known as the *Hadamard conjecture*, although it is not in Hadamard's papers; it was first explicitly stated by Paley.

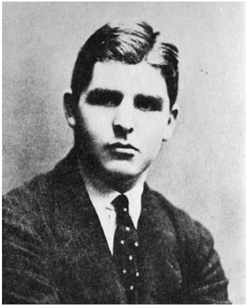
Paley (1933) showed how to construct a Hadamard matrix of order $q + 1$ when $q \equiv 3 \bmod 4$ is a prime power, and of order $2(q + 1)$ when $q \equiv 1 \bmod 4$ is a prime power. Combined with a doubling construction of Sylvester (1867), this shows that Hadamard matrices of order $n = 2^r(q + 1)$ exist whenever q is zero or an odd prime power, $r \geq 0$ and $4|n$.

Many other constructions have been found. Since 2005 it has been known that all $n = 4k \leq 664$ are the orders of Hadamard matrices. However, it is not known if the Hadamard orders have a positive density in $\mathbb{N}$ (compare the sequence of primes).

The Hadamard conjecture (Paley's conjecture)

> *It seems probable that, whenever n is divisible by 4, it is possible to construct an orthogonal matrix of order n composed of ± 1, but the general theorem has every appearance of difficulty.*

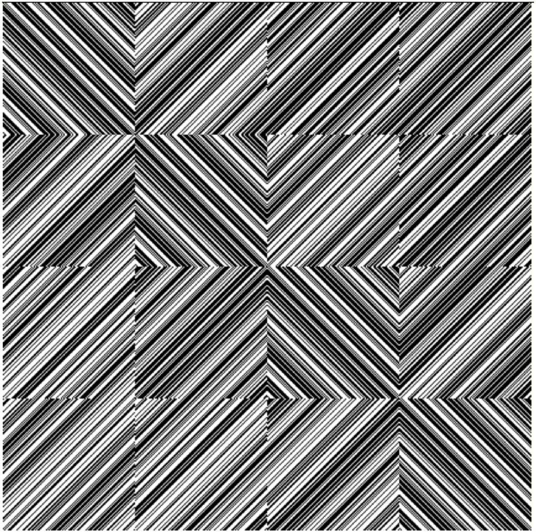
Paley, 1933



Raymond Paley (1907–1933)

Paley was killed by an avalanche while skiing near Banff in the Canadian Rockies.

A Hadamard matrix of order 428



A Hadamard matrix of order 428 constructed by Kharaghani and Tayfeh-Rezaie (2005); since then 668 has been the smallest order $n = 4k$ for which a construction (or existence proof) is not known.

$D(n)$ 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.

Recall that n is a *Hadamard order* if a Hadamard matrix of order n exists, and a *non-Hadamard order* otherwise.

For example, $1, 2, 4, 8, 12, 16, 20, 24$ are Hadamard orders; $3, 5, 6, 7, 9, 10, 11, 13$ are non-Hadamard orders.

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

$R(n)$ 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 $n \le 31$ ("?" means conjectured)

Each block of two columns corresponds to a congruence class of $n \bmod 4$. Within the columns of $R(n)$ values there are interesting oscillations. Data from Will Orrick's website `http://www.indiana.edu/~maxdet/`.

The Barba bound

For the three congruence classes $n \equiv 1, 2, 3 \bmod 4$ there are specialised upper bounds on $R(n)$ that are slightly sharper than the Hadamard bound $R(n) \le 1$.

For example, if $n \equiv 1 \bmod 4$, there is an upper bound due to Barba (1933):

$$R(n) \le (2n-1)^{1/2}(n-1)^{(n-1)/2}/n^{n/2} \sim (2/e)^{1/2} \approx 0.86.$$

This bound is attained for all $n = q^2 + (q+1)^2$, where q is an odd prime power [Brouwer, 1983], as well as in the small cases $q \in \{1, 2, 4\}$.

Two strategies for lower bounds

There are two ways that we can obtain a lower bound on $D(n)$ or $R(n)$ 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. There are theorems about minors of Hadamard matrices which give 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.

We consider bordering as it gives better results in general, and the probabilistic method is applicable to it.

Conjectured lower bound on $R(n)$

It appears plausible that there always exists such a matrix with determinant greater than $\frac{1}{2}h_n$*, where* $h_n = n^{n/2}$ *is the Hadamard bound.*

Rokicki, Orrick et al (2010)

Tomas Rokicki

Will Orrick

Known lower bounds on $R(n)$

What can we say about lower bounds on $R(n)$?

Rokicki, Kazmenko, Meyrignac, Orrick, Trofimov and Wroblewski (2010) verified numerically that $R(n) > 1/2$ for all $n \le 120$, and conjectured that this lower bound always holds.

However, the theoretical bounds are much weaker.

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

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

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 $d = n - h$.

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

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

The bound $\frac{1}{9}$ is independent of n (and does not tend to zero).

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

A naive approach

How can we use the probabilistic method to give a lower bound on $R(n)$?

An obvious approach is to consider a random $\{\pm 1\}$-matrix of order n, hoping that a random matrix often has a large determinant. (It does, but not large enough!)

In 1940, Turán showed that the

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

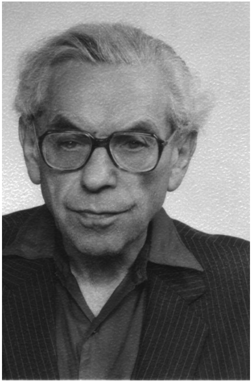
for $\{\pm 1\}$-matrices A of order n, chosen uniformly at random.

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

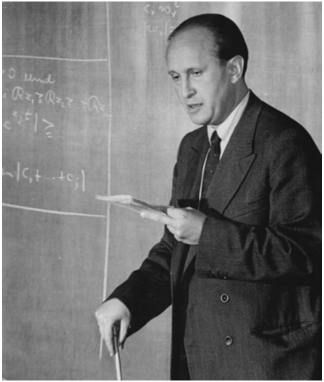
$$\mathbb{E}[\det(A)^2] = n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \ll n^n.$$

This weaker than what we need by a factor of almost e^n.

Erdős and Turán



Pál Erdős (1913–1996)



Pál Turán (1910–1976)

Chebyshev's inequality

We use Chebyshev's classical "tail inequality".

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 $\lambda > 0$,

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

For example, 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). Let's take $\lambda = n^{n/2}/2$ (half the Hadamard bound). Then

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

is tiny if n is large. Thus, large-determinant matrices are rare!

A different approach – bordering a Hadamard matrix

Suppose $n = h + d$ where h is the order of a Hadamard matrix H, and d is small. (If the Hadamard conjecture is true, we can assume that $0 \le d \le 3$.)

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 order n matrix will often have a large determinant.

To analyse the effect of a border on the determinant, we need to look at 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 $h \times h$, and $n = h + d > h$. (Here H does not have to be Hadamard, any nonsingular $h \times h$ matrix will do.)

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

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

The Schur complement is relevant to our problem because

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

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

Proof of the determinant identity

To prove the Schur complement identity

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

just take determinants of each side in the identity

$$A = \begin{bmatrix} H & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CH^{-1} & I \end{bmatrix} \begin{bmatrix} H & B \\ 0 & D - CH^{-1}B \end{bmatrix} .$$

You can verify this "block LU factorization" directly by block matrix multiplication, or derive it by block Gaussian elimination.

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, where $n = h + d$ as usual.

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

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

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

- The problem is to maximise the order d determinant

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

(The sign of the determinant is not important, only the absolute value is of interest to us.)

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

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

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 $C \cdot H^T B$, then choose D to maximise $|\det|$. 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: a good construction

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

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

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

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

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

Finally, choose $D = -I$ (we can adjust the off-diagonal elements of D later).

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]

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

- 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

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

but

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

What went wrong?

The problem 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

Looking at 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 if (and only if)

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

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

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

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

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

Using these two inequalities and the fact that the diagonal elements of F are independent, we can get a useful lower bound on $\mathbb{E}[\det(F)]$. I will show you the cases $d = 2, 3$.

In both cases we get the correct order of magnitude, $h^{d/2}$.

Lower bound on $\mathbb{E}[\det(F)]$ when $d = 2$

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

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

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

where $O(1)$ means some constant, independent of h.

Lower bound on $\mathbb{E}[\det(F)]$ when $d = 3$

If $d = 3$, a similar argument is

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

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

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

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

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

Lower bounds for $d \le 3$

Using the fact that there must exist a matrix F_0 such that $\det(F_0) \ge \mathbb{E}[\det(F)]$, and explicitly bounding the error terms, we[3] can prove:

Theorem. If $0 \le d \le 3$, $n = h + d$, where h is the order of a Hadamard matrix, then

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

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.

[3] Brent, Osborn and Smith, arXiv:1501.06235v1.

Lower bound for arbitrary d

If we don't assume the Hadamard conjecture, then $d > 3$ is possible. How large can d be?

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

Unfortunately, the argument that we used for $d \le 3$ involves expanding $\det(F)$ to give $d!$ terms. We then approximate $\det(F)$ by the "diagonal" term $f_{11}f_{22}\cdots f_{dd}$ and bound the contribution of the remaining $(d! - 1)$ terms.

The main term is of order $h^{d/2}$ and the sum the other terms is of order $d!h^{d/2-1}$. Thus, this approach is only useful when

$$h \gg d!$$

From Livinskyi's result, we can assume that $h \gg d^6$, but this is not large enough. Hence, we need a different approach.

Ostrowski's inequality

Chebyshev's inequality and a theorem of Ostrowski allow us to avoid an expansion involving $d!$ terms.

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 the 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, and the off-diagonal elements have mean zero and variance 1.

Livinskyi, Ostrowski, Chebyshev

Ivan Livinskyi

Alexander Ostrowski (1893–1986)

Pafnuty Chebyshev (1821–1894)

Digression

One of the great things about being one of Gene's students at Stanford was that there were so many visitors who came to work with Gene and/or give seminars. Sometimes they even stayed for several months and taught a course. In this way I was lucky enough to meet Ostrowski as well as Björck, Bunch, Dahlquist, Dongarra, Duff, Gear, Henrici, Kahan, Moler, Parlett, Stewart, Varga, Wilkinson, . . .

The choice of D

We can no longer ignore the bottom right $d \times d$ matrix 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.

This 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 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 (the opposite of 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 $h \geq 4$,

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

Thus, we can take $\sigma^2 \leq 1/4$ in Chebyshev's inequality. This gives

$$\mathbb{P}[g_{ii} \text{ is bad}] \leq \frac{\sigma^2}{d^2} \leq \frac{1}{4d^2}\,.$$

Thus, the probability that *any* diagonal element is bad is at most $1/(4d)$.

Putting the pieces together

Putting the pieces together,

$$\mathbb{P}[G \text{ is bad}] \le \left(1 - \frac{1}{d}\right) + \frac{1}{4d} < 1.$$

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, *some* choice of B must give a good G.

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.

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

This leads to the following theorem, which gives a useful inequality provided $d^2 < \mu$.

Theorem. If $d \geq 1$, $n = h + d$ there exists a Hadamard matrix H of order h, and $\mu \sim \sqrt{2h/\pi}$ is as above, 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 the condition $d^2 < \mu$ is satisfied for all sufficiently large n.

The lower bounds on $D(n)$ and $R(n)$

Theorem again. If $d \geq 1$, $n = h + d$ there exists a Hadamard matrix H of order h, and $\mu \sim \sqrt{2h/\pi}$ is as above, then

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

In this lower bound, 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 Ostrowski's inequality.

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 $d^2/h^{1/2} \to 0$ as $n \to \infty$, this is close to the bound $R(n) \geq (2/(\pi e))^{d/2}$ that we obtained for $d \leq 3$.

Randomised algorithms

The probabilistic construction can easily be used to give a randomised algorithm for finding large-determinant matrices, i.e. nearly D-optimal designs.

The algorithm actually works better than the theory suggests. In all the cases that we have tried, it is easy to find an $n \times n$ $\{\pm 1\}$-matrix A with

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

In practice, the main difficulty is in constructing a Hadamard matrix of the required order h, because constructions for Hadamard matrices are scattered throughout the literature and sometimes appeared in obscure journals or conference proceedings. There is no known "efficient" and "uniform" way to construct a Hadamard matrix of given order h (if it exists).

We do not want to try all 2^{h^2} possibilities!

Optimality of the bounds

Consider the inequality

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

which our probabilistic algorithm suggests is always true (and is close to what we can prove).

The factor $(2/e)^{d/2}$ is asymptotically optimal (as $n \to \infty$) for $d \leq 2$; we do not know if it is asymptotically optimal for $d \geq 3$.

The factor $\pi^{-d/2}$ is not optimal, but seems to be an inherent limitation imposed by the probabilistic method, which is estimating a mean rather than a maximum.

Conclusion

We've seen that probabilistic ideas are useful for

- proving lower bounds close to Hadamard's upper bound on the largest-possible determinants of $\{\pm 1\}$-matrices of a given order, and
- finding large-determinant $\{\pm 1\}$-matrices (optimal designs).

I hope you are convinced that probabilistic ideas are relevant even for problems that do not appear to involve any randomness. There are many other examples that I could have given if we had more time. See, for example, the book by Alon and Spencer.

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:1501.06235v1, 26 Jan. 2015.

R. W. Cottle, Manifestations of the Schur complement, *Linear Alg. Appl.* 8 (1974), 189–211.

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

G. H. Golub and C. F. Van Loan, Matrix computations, 3rd edn., Johns Hopkins Press, 1996.

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

H. Kharaghani and B. Tayfeh-Rezaie, A Hadamard matrix of order 428, *J. Combin. Designs* 13 (2005), 435–440.

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`

Wikipedia, Optimal design,
`http://en.wikipedia.org/wiki/Optimal_design`.