UPPER BOUNDS FOR THE SPECTRAL RADIUS OF THE $n \times n$ HILBERT MATRIX

PETER OTTE
UPPER BOUNDS FOR THE SPECTRAL RADIUS
OF THE $n \times n$ HILBERT MATRIX

PETER OTTE

We derive upper bounds for the spectral radius of the $n \times n$ Hilbert matrix.
The key idea is to write the Hilbert matrix as integral operator with positive
kernel function and then to use a Wielandt-type min-max principle for the
spectral radius. Choosing special trial functions yields a new bound that
improves the best bound known heretofore.

1. Introduction

The spectral asymptotics of the Hilbert matrix has attracted a lot of interest concerning
both the lowest and the largest eigenvalue. Here we shall focus on the spectral
radius $\rho_n$ of the $n \times n$ Hilbert matrix for which we shall prove, particularly, the
bound

\[
\rho_n \leq 2 w_n \arcsin \frac{1}{w_n} \quad \text{with} \quad w_n := 2 \left( \frac{(n!)^2}{(2n)!} \right)^{1/2n}, \quad n \in \mathbb{N}.
\]

This improves, at least for large values of $n$, Cassels’ bound, given in (5) below,
which is the best hitherto known. Numerical computations suggest that (1) is ac-
tually better for all $n$ except $n = 1, 2$.

We base the proof of (1) upon relating the Hilbert matrix to an integral operator
$H_n$ whose spectral radius can be expressed by a min-max principle for operators
having positive kernel functions:

\[
\rho_n = \inf_{\varphi \in M} \sup_{0 < x < 1} \frac{(H_n \varphi)(x)}{\varphi(x)},
\]

where $M$ is some set of appropriate trial functions. For the sake of completeness
we shall prove (2) without recourse to the general theory. In the matrix case the
above min-max principle is due to Wielandt [1950] and related to the enclosure
result of Collatz [1942]. It has been generalized in many directions; see [Friedland
1990; Marek 1966; Schaefer 1984], for example.


Keywords: Hilbert matrix, Hilbert inequality, spectral radius, Wielandt min-max principle, integral
operator.
To derive estimates we pick $\varphi(x) := (1 - x)^{\gamma}$ in (2), with $-1 < \gamma < 0$. We restrict ourselves to the case $\gamma = -\frac{1}{2}$, for which the calculations are manageable, and obtain (1).

Hilbert was the first one to investigate spectral properties of the matrix named after him. In his lectures he showed his double series theorem stating that $\rho_n$ stays finite as $n \to \infty$; this was first published by Weyl [1908] (see also [Wiener 1910]). The concrete inequality

$$\rho_n \leq \pi$$

is due to Schur [1911]. This is the optimal constant that does not depend on the dimension $n$. However, if we do want the bound to depend on $n$ it is possible to strengthen (3). Frazer [1946] obtained

$$\rho_n \leq n \sin \frac{\pi}{n} \quad \text{for } n \geq 2,$$

by refining a method of Fejér and Riesz [1921], which they used to prove what is now called the Fejér–Riesz inequality for analytic functions. Equation (4) was later rediscovered by Hsiang [1957] and Yahya [1965], and was eventually improved by Cassels [1948] to

$$\rho_n \leq 2 \arctan \sqrt{2n}.$$  

Finally, it might be instructive to look at the asymptotic expansion of $\rho_n$. The first asymptotic result was obtained by Taussky [1949] by computing the quadratic form with special trial vectors having components $c_k := 1/\sqrt{k}$; it was

$$\rho_n = \pi + O\left(\frac{1}{\ln n}\right).$$

The exact asymptotic behaviour

$$\rho_n = \pi - \frac{\pi^5}{2 \ln^2 n} + O\left(\frac{\ln \ln n}{\ln^3 n}\right)$$

was determined by de Bruijn and Wilf [1962], who compared the matrix operator with an integral operator whose spectral asymptotics can be derived from general results of Widom [1958] (see also [Widom 1961]).

2. Estimates for the spectral radius

We start by relating the Hilbert matrix

$$A_n := \left(\frac{1}{j+k+1}\right)_{j,k=0,...,n-1}$$

(6)
to the integral operator \( H_n : C[0, 1] \to C[0, 1] \) having the kernel function

\[
K_n(xy) := \sum_{j=0}^{n-1} (xy)^j = \frac{1 - (xy)^n}{1 - xy}.
\]

For \( n = \infty \) this operator was used by Magnus [1950] to study the spectrum of the infinite Hilbert matrix. We let \( H_n \) act on \( C[0, 1] \) because we want to have sufficiently many trial functions at hand. As we hoped, \( H_n \) has (almost) the same spectrum as \( A_n \). In particular, they have the same spectral radius, henceforth denoted by \( \rho_n \).

**Lemma 1.** Let \( C[0, 1] \) be equipped with the usual maximum norm. Then \( H_n : C[0, 1] \to C[0, 1] \) is a bounded linear operator. The respective spectra of the Hilbert matrix \( A_n \) and the integral operator \( H_n \) are the same apart from 0. Their common spectral radius \( \rho_n \) can be expressed by

\[
\rho_n = \inf_{\varphi \in M} \sup_{0 < x < 1} \frac{(H_n \varphi)(x)}{\varphi(x)}, \quad \text{where} \quad M := \{ \varphi \in L^1[0, 1] \mid \varphi > 0, \frac{1}{\varphi} \in C[0, 1] \}.
\]

**Proof.** It is clear from the definition and (7) that \( H_n \) is linear and bounded. Also (7) shows that \( H_n \) has \( n \)-dimensional range spanned by the monomials \( x^k \), for \( k = 0, \ldots, n - 1 \), which implies that the spectrum of \( H_n \) consists only of eigenvalues. To each \( c \in \mathbb{C}^n \) we associate \( \varphi_c \in C[0, 1] \) in the natural way:

\[
c = (c_0, \ldots, c_{n-1}) \in \mathbb{C}^n \quad \longleftrightarrow \quad \varphi_c(x) = \sum_{j=0}^{n-1} c_j x^j.
\]

The statement on the spectra then follows from

\[
(H_n \varphi_c)(x) = \int_0^1 \left( \sum_{j=0}^{n-1} (xy)^j \right) \sum_{k=0}^{n-1} c_k y^k \, dy
\]

\[
= \sum_{j,k=0}^{n-1} c_k x^j \int_0^1 y^{j+k} \, dy = \sum_{j=0}^{n-1} x^j \sum_{k=0}^{n-1} \frac{1}{j+k+1} c_k.
\]

Note that \( H_n \) must have a kernel and \( A_n \) does not.

To prove Formula (8) we recall from the Perron–Frobenius Theorem that, since \( A_n \) has positive entries, \( \rho_n \) is an eigenvalue of \( A_n \) and hence of \( H_n \). Let \( v \) be the corresponding eigenfunction. Writing down the eigenvalue equation for \( v \) and dividing by \( \varphi \in M \) yields

\[
\rho_n \frac{v(x)}{\varphi(x)} = \int_0^1 K_n(xy) \frac{\varphi(y)}{\varphi(x)} \frac{v(y)}{\varphi(y)} \, dy.
\]
This shows that \( v/\varphi \in C[0, 1] \) is an eigenfunction of the operator \( H_{n, \varphi} \) with kernel
\[
K_{n, \varphi}(xy) := K_n(xy) \frac{\varphi(y)}{\varphi(x)},
\]
whence \( \rho_n \leq \rho(H_{n, \varphi}) \), the spectral radius of \( H_{n, \varphi} \). Since \( \rho(H_{n, \varphi}) \leq \|H_{n, \varphi}\|_{\infty} \) we conclude that
\[
\rho_n \leq \|H_{n, \varphi}\|_{\infty} = \sup_{0 < x < 1} \int_0^1 K_{n, \varphi}(xy) \, dy
\]
where we have used \( \varphi(x) > 0 \), \( K_n(xy) \geq 0 \), and thus \( K_{n, \varphi}(xy) \geq 0 \). To show equality in (8) we once again invoke the Perron–Frobenius Theorem, according to which the eigenvector of \( A_n \) belonging to \( \rho_n \) can be chosen to have positive components, whence we can, via (9), likewise choose the eigenfunction \( v > 0 \). In particular, \( v \in M \). \( \square \)

We use Lemma 1 to estimate the spectral radius from above by cleverly choosing trial functions in (8):

\[
r_n(x) := \frac{(H_n \varphi)(x)}{\varphi(x)} = \frac{1}{\varphi(x)} \int_0^1 \frac{1 - (xy)^n}{1 - xy} \varphi(y) \, dy \quad \text{for } n \in \mathbb{N}.
\]

To get an idea of what the \( \varphi \)'s should look like we cast \( r_n \) into a form more amenable to further investigation. The crucial point is to evaluate the integral
\[
J_n(x) := \int_0^1 \frac{y^n}{1 - xy} \varphi(y) \, dy.
\]
We start by differentiating with respect to \( x \):

\[
J_n'(x) = \int_0^1 \frac{y^{n+1}}{(1 - xy)^2} \varphi(y) \, dy = \frac{1}{x} \int_0^1 \frac{y^n}{(1 - xy)^2} \varphi(y) \, dy - \frac{1}{x} J_n(x).
\]

The explicitly written integral on the right can also be produced by integration by parts, which we perform in such a way that \( \varphi(1) \) is omitted because our trial functions will have a singularity at \( x = 1 \):

\[
J_n(x) = \left[ (y - 1) \frac{y^n}{1 - xy} \varphi(y) \right]_0^1 - \int_0^1 (y - 1) \left( \frac{ny^{n-1}}{1 - xy} + \frac{xy^n}{(1 - xy)^2} \right) \varphi(y) + \frac{y^n}{1 - xy} \varphi'(y) \, dy
\]
\[
= \delta_n \varphi(0) + n J_{n-1}(x) - n J_n(x) + (x - 1) \int_0^1 \frac{y^n}{(1 - xy)^2} \varphi(y) \, dy
\]
\[
+ J_n(x) + \tilde{J}_n(x),
\]

where we have used \( \varphi(x) > 0 \), \( K_n(xy) \geq 0 \), and thus \( K_{n, \varphi}(xy) \geq 0 \).
with \( \delta_n := \delta_{n,0} \) the Kronecker delta and

\[
\tilde{J}_n(x) := \int_0^1 \frac{y^n}{1-xy}(1-y)\varphi'(y) \, dy.
\]

Hence we can eliminate the integral in question from (11):

\[
(12) \quad J'_n(x) = \frac{1}{x(1-x)} \left( \delta_n \varphi(0) + n J_{n-1}(x) - n J_n(x) + \tilde{J}_n(x) \right) - \frac{1}{x} J_n(x).
\]

To eliminate the annoying \( J_{n-1} \) we observe that

\[
J_n(x) = \int_0^1 \frac{y^n}{1-xy} \varphi(y) \, dy = \frac{1}{x} J_{n-1}(x) - \frac{1}{x} \int_0^1 y^{n-1} \varphi(y) \, dy,
\]

and therewith (12) becomes

\[
(13) \quad J'_n(x) = \frac{\delta_n}{x(1-x)} \varphi(0) - \frac{n+1}{x} J_n(x) + \frac{\kappa_n}{x(1-x)} + \frac{1}{x(1-x)} \tilde{J}_n(x)
\]

where we have put

\[
(14) \quad \kappa_0 := 0, \quad \kappa_n := n \int_0^1 y^{n-1} \varphi(y) \, dy \quad \text{for} \quad n \in \mathbb{N}.
\]

We are going to express

\[
(15) \quad \Phi_n(x) := \frac{x^n}{\varphi(x)} J_n(x)
\]

by dint of (13) through a differential equation:

\[
\Phi'_n(x) = - \frac{\varphi'(x)}{\varphi^2(x)} x^n J_n(x) + \frac{n x^{n-1}}{\varphi(x)} J_n(x) + \frac{x^n}{\varphi(x)} J'_n(x)
\]

\[
= - \left( \frac{\varphi'(x)}{\varphi(x)} + \frac{1}{x} \right) \Phi_n(x) + \frac{x^{n-1}}{(1-x)\varphi(x)} \tilde{J}_n(x) + \frac{x^{n-1}}{(1-x)\varphi(x)} (\delta_n \varphi(0) + \kappa_n).
\]

At this point we fix our trial function \( \varphi \) in such a way that

\[
(16) \quad (1-x)\varphi'(x) = -\gamma \varphi(x),
\]

that is, \( \varphi(x) = (1-x)^\gamma \) with some \( \gamma \in \mathbb{R} \) to be specified later, whereby \( \tilde{J}_n \) becomes a multiple of \( J_n \), and we arrive at a differential equation for \( \Phi_n \):

\[
(17) \quad \Phi'_n(x) = - (\gamma + 1) - \frac{1}{x} \Phi_n(x) + \frac{x^{n-1}}{(1-x)^{1+\gamma}} (\delta_n + \kappa_n).
\]

This is equivalent to

\[
(x^{1+\gamma} \Phi_n(x))' = \frac{x^{n+\gamma}}{(1-x)^{1+\gamma}} (\delta_n + \kappa_n),
\]
which we can solve immediately for $\Phi_n$:

\begin{equation}
\Phi_n(x) = \frac{\delta_n + \kappa_n}{x^{1+\gamma}} \int_0^x \frac{\xi^{n+\gamma}}{(1-\xi)^{1+\gamma}} \, d\xi \quad \text{for } n \in \mathbb{N}_0.
\end{equation}

In particular, we can now see that $\gamma$ must satisfy $-1 < \gamma < 0$ in order to yield well-defined integrals and to have $\varphi \in M$ in (8). We summarize our calculations.

**Theorem 2.** The spectral radius $\rho_n$ of the $n \times n$ Hilbert matrix $A_n$ as in (6) can be estimated by

\begin{equation}
\rho_n \leq \inf_{0<\alpha<1} \sup_{0<x<1} \frac{1}{x^{1-\alpha}} \int_0^x \frac{x - \kappa_n x^n}{\xi^\alpha (1-\xi)^{1-\alpha}} \, d\xi,
\end{equation}

where

\begin{equation}
\kappa_n = \frac{n!}{(n-\alpha)(n-1-\alpha) \cdots (1-\alpha)} \quad \text{for } n \in \mathbb{N}.
\end{equation}

**Proof.** Put $\alpha := -\gamma$ and use in turn the min-max principle (8) and the definitions of $r_n$ and $\Phi_n$ as in (10) and (15), with $\varphi$ being chosen according to (16) to obtain

$$
\rho_n \leq \inf_{0<\alpha<1} \sup_{0<x<1} r_n(x) = \inf_{0<\alpha<1} \sup_{0<x<1} (\Phi_0(x) - \Phi_n(x)).
$$

Then (19) follows directly from the representation (18) of $\Phi_n$.

For $\varphi$ as in (16) the integral in (14) is Euler's beta function. Hence,

$$
\kappa_n = n B(n, 1-\alpha) = \frac{n \Gamma(n) \Gamma(1-\alpha)}{\Gamma(n+1-\alpha)}
$$

wherefrom we deduce (20). \qed

The optimal way to derive bounds on $\rho_n$ would be to determine the maximum of the function $r_n$ exactly. Unfortunately, this turns out to be rather complicated, and we content ourselves with narrowing the region where the maximum must lie.

**Corollary 3.** The spectral radius $\rho_n$ of the $n \times n$ Hilbert matrix $A_n$ can be estimated by

\begin{equation}
\rho_n \leq \inf_{0<\alpha<1} \kappa_n^{(1-\alpha)/n} \int_0^{1/\kappa_n^{1/n}} \frac{1}{\xi^\alpha (1-\xi)^{1-\alpha}} \, d\xi,
\end{equation}

which in the case $\alpha = \frac{1}{2}$ specializes to

\begin{equation}
\rho_n \leq 2 w_n \arcsin \frac{1}{w_n} \quad \text{with } w_n := \kappa_n^{1/2n} = 2 \left( \frac{(n!)^2}{(2n)!} \right)^{1/2n}.
\end{equation}
Proof. When \( 1 - \kappa_n \xi^n \leq 0 \) the function \( r_n \) is decreasing, whence the maximum must lie in the interval \([0, x_0]\) for \( x_0 \) the unique zero of the integrand in (19):

\[
1 - \kappa_n x_0^n = 0, \quad \text{i.e.,} \quad x_0 = 1/\kappa_n^{1/n}.
\]

We conclude

\[
\sup_{0 < x < 1} r_n(x) = \sup_{0 < x \leq x_0} r_n(x) \leq \sup_{0 < x \leq x_0} \Phi_0(x) = \Phi_0(x_0)
\]

because the \( \Phi_n(x) \) are nonnegative and \( \Phi_0 \) increases.

For \( \alpha = \frac{1}{2} \) the \( w_n \) are easily obtained from (22) and (20) and the integral in (21) can be evaluated by the change of variables \( \xi = s^2 \).

Finally, we shall check that our estimate (22) is indeed better than (5). Using some familiar formulae for \( \arctan \) and \( \arcsin \) we obtain

\[
\arctan \sqrt{2n} - w_n \arcsin \frac{1}{w_n} = \arctan \sqrt{2n} - \arcsin \frac{1}{w_n} - (w_n - 1) \arcsin \frac{1}{w_n}
\]

\[
\geq \arctan \sqrt{2n} - \arctan \frac{1}{\sqrt{w_n^2 - 1}} - \frac{\pi}{2} (w_n - 1)
\]

\[
= \arctan \frac{\sqrt{2n(w_n^2 - 1)} - 1}{\sqrt{w_n^2 - 1 + 2n}} - \frac{\pi}{2} (w_n - 1).
\]

Now the asymptotics of the middle binomial coefficient yields

\[
w_n \sim 2 \left( \frac{\sqrt{\pi n}}{4^n} \right)^{1/2n} \sim n^{1/4n} \quad \text{as} \quad n \to \infty,
\]

which implies immediately \( \lim_{n \to \infty} w_n = 1 \), and further

\[
n(w_n - 1) \sim n(n^{1/4n} - 1) = n(e^{(\ln n)/4n} - 1) \sim \frac{1}{4} \ln n \quad \text{as} \quad n \to \infty.
\]

Therefore, for large \( n \),

\[
\frac{\sqrt{2n(w_n^2 - 1)} - 1}{\sqrt{w_n^2 - 1 + 2n}} \geq \frac{1}{4} \sqrt{2n(w_n^2 - 1)} \geq \frac{1}{4} \sqrt{w_n - 1}.
\]

Noting \( \arctan x \geq cx \) for small \( x \) with some constant \( c > 0 \) and using the monotonicity of \( \arctan \) we conclude

\[
\arctan \sqrt{2n} - w_n \arcsin \frac{1}{w_n} \geq \frac{c}{4} \sqrt{w_n - 1} - \frac{\pi}{2} (w_n - 1) > 0
\]

for large values of \( n \). With some care it should be possible to show the statement for smaller values of \( n \), too.
3. Remarks

We suggest some topics that might be worth further investigation.

(1) In order to derive from Theorem 2 a bound that can be computed more or less explicitly we did not determine in (19) the maximum of the function $r_n$ exactly. Thus, the first possibility to strengthen (22) is to study the maximum of $r_n$.

(2) Also for computational reasons we fixed the exponent $\alpha = \frac{1}{2}$ in (21). However, numerical computations suggest that $\alpha = \frac{1}{2}$ is generally not the optimal choice and that other values of $\alpha$ give much more accurate estimates. According to a theorem of Čebyšev the integral in (21) can be evaluated for any $0 < \alpha < 1$ in closed form by means of elementary functions. It is not clear whether these elementary functions allow for an efficient minimizing procedure.

(3) A vaguer idea is to pick other trial functions than $(1 - x)^{-\alpha}$. Our method will work as long as we arrive at a differential equation for $\Phi_n$ as in (17).

(4) Since Wielandt’s min-max principle is accompanied by a max-min principle, one can also think of deriving lower bounds for the spectral radius in which case; however, completely different trial functions are needed.

Acknowledgments

I thank Robert Denk and Hubert Kalf for stimulating discussions and their steady interest in this subject.

References


PETER OTTE
RUHR-UNIVERSITÄT BOCHUM
FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄTSSTRASSE 150
D-44780 BOCHUM
GERMANY
peter.otte@ruhr-uni-bochum.de