ESTIMATES OF THE DETERMINANT OF A PERTURBED IDENTITY MATRIX*

SIEGFRIED M. RUMP †

Abstract. Recently Brent et al. presented new estimates for the determinant of a real perturbation $I + E$ of the identity matrix. They give a lower and an upper bound depending on the maximum absolute value of the diagonal and the off-diagonal elements of E, and show that either bound is sharp. Their bounds will always include 1, and the difference of the bounds is at least $\mathrm{tr}(E)$. In this note we present a lower and an upper bound depending on the trace and Frobenius norm $\epsilon := \|E\|_F$ of the (real or complex) perturbation E, where the difference of the bounds is not larger than $\epsilon^2 + \mathcal{O}(\epsilon^3)$ provided that $\epsilon < 1$. Moreover, we prove a bound on the relative error between $\det(I + E)$ and $\exp(\mathrm{tr}(E))$ of order ϵ^2.

Key words. Determinant, Hadamard bound, Ostrowski bound, Hans-Schneider bound, perturbation of identity

AMS subject classifications. 15A15, 65F40

1. Introduction and main results. Classical estimates for the determinant of a matrix include the Hadamard bound [7] or Gershgorin circles [6]. Moreover, Ostrowski [11, 12, 13] gave a number of lower and upper bounds. Other estimates include [4, 9, 1]. In particular, bounds for the determinant of a perturbed identity matrix are given in Ostrowski's papers, or in [15].

Recently, new sharp bounds for $\det(I + E)$ have been presented by Brent et al. in [2, 3]. Denote by δ the maximum absolute value of the diagonal elements, and by ε the maximum absolute value of the off-diagonal elements of a real $n \times n$-matrix E. Then [2, 3] prove

$$(1 - \delta - (n-1)\varepsilon)(1 - \delta + \varepsilon)^{n-1} \le \det(I + E) \le ((1+\delta)^2 + (n-1)\varepsilon^2)^{n/2}, \tag{1}$$

where $\delta + (n-1)\varepsilon \le 1$ is supposed for the left inequality. Both inequalities are sharp as by explicit examples given in [2, 3]. For convergent E, Fredholm's identity [5]

$$\det(I + E) = \exp\left(\sum_{k=1}^{\infty} (-1)^{k-1} \frac{\mathrm{tr}(E^k)}{k}\right) \tag{2}$$

yields $\det(I + E) = \exp(\mathrm{tr}(E)) + \mathcal{O}(\varepsilon^2)$ for $\|E\| \le \varepsilon < 1$ and some matrix norm $\|\cdot\|$. This is reflected in (1). Although being individually sharp, the upper and lower bound in (1) always include the number 1 and differ by at least $\mathrm{tr}(E)$. That is also true for most of the other bounds mentioned.

Notable exceptions are papers by Ostrowski [14] and Hans Schneider [16], proving bounds depending on the trace and on the absolute row sums of E. If all elements of E are bounded by ε in absolute value, then either difference between upper and lower bound is $\mathcal{O}(n^3\varepsilon^2)$.

For real or complex E, we prove two-sided bounds differing by $\mathcal{O}(\epsilon^2)$, where $\epsilon := \|E\|_F = [\mathrm{tr}(E^H E)]^{1/2}$ denotes the Frobenius (or Hilbert-Schmidt) norm. We prove absolute bounds on $|\det(I + E)|$, and relative bounds on $\det(I + E)$.

*This research was partially supported by CREST, Japan Science and Technology Agency.

†Institute for Reliable Computing, Hamburg University of Technology, Am Schwarzenberg-Campus 3, Hamburg 21073, Germany, and Visiting Professor at Waseda University, Faculty of Science and Engineering, 3–4–1 Okubo, Shinjuku-ku, Tokyo 169–8555, Japan (rump@tuhh.de).

THEOREM 1.1. *Let E be a real or complex $n \times n$ matrix. Then*

$$|\det(I+E)| \leq \exp\left(\Re(\operatorname{tr}(E)) + \frac{\epsilon^2}{2}\right). \tag{3}$$

Suppose the eigenvalues λ_k of E satisfy $\Re(\lambda_k) > -1$, and denote $\mu_k := \min(0, \Re(\lambda_k))$. Then

$$\exp\left(\Re(\operatorname{tr}(E)) - \frac{\epsilon^2/2}{1+\min\limits_k \mu_k}\right) \leq |\det(I+E)|. \tag{4}$$

Denote the spectral radius by $\rho(\cdot)$. If $\rho(E) < 1$, then

$$\exp\left(\Re(\operatorname{tr}(E)) - \frac{\epsilon^2/2}{1-\rho(E)}\right) \leq |\det(I+E)|. \tag{5}$$

If $\epsilon < 1$, then

$$\exp\left(\Re(\operatorname{tr}(E)) - \frac{\epsilon^2}{2(1-\epsilon)}\right) \leq |\det(I+E)| \leq \exp\left(\Re(\operatorname{tr}(E)) + \frac{\epsilon^2}{2}\right), \tag{6}$$

The denominator in the lower bound of (6) *cannot be replaced by* 2.

REMARK 1. Note that $\Re(\lambda_k) > -1$ implies $\det(I+E) = |\det(I+E)|$ for real E.

REMARK 2. Computationally, an upper bound on $\rho(E)$ is easily obtained by Perron-Frobenius Theory and $\rho(E) \leq \rho(|E|) \leq \max_i \frac{(|E|x)_i}{x_i}$ for any positive vector x, with the Perron vector of $|E|$ being optimal.

REMARK 3. The upper bound in (3) is given to show the symmetry to the following lower bounds; it is never better than Hadamard's bound:

$$|\det(I+E)| \leq \prod_{k=1}^{n} \|(I+E)_{k*}\|_2 \leq \exp\left(\Re(\operatorname{tr}(E)) + \frac{\epsilon^2}{2}\right), \tag{7}$$

where M_{k*} denotes the k-th row of a matrix M.

THEOREM 1.2. *Let E be a real or complex $n \times n$ matrix and suppose $\rho(E) < 1 - \epsilon^2/2$. Then*

$$\left|\frac{\det(I+E) - \exp(\operatorname{tr}(E))}{\exp(\operatorname{tr}(E))}\right| \leq \frac{\epsilon^2}{2(1-\rho(E)-\epsilon^2/2)} \leq \frac{\epsilon^2}{2(1-\epsilon-\epsilon^2/2)}. \tag{8}$$

Except for E being the zero matrix, the implied upper bound on $|\det(I+E)|$ is always worse than Hadamard's bound:

$$\prod_{k=1}^{n} \|(I+E)_{k*}\|_2 \leq |\exp(\operatorname{tr}(E))| \left(1 + \frac{\epsilon^2}{2(1-\epsilon^2/2)}\right). \tag{9}$$

If $\epsilon \leq 0.5173$, then

$$\left|\frac{\det(I+E) - \exp(\operatorname{tr}(E))}{\exp(\operatorname{tr}(E))}\right| \leq \frac{\epsilon^2}{2(1-\epsilon)}. \tag{10}$$

The denominator in the bound cannot be replaced by 2.

2. Proofs. We need the following facts. Let E be a real or complex $n \times n$ matrix with eigenvalues λ_k. Then

$$|\det(I+E)| = \exp\left(\frac{1}{2}\sum_{k=1}^{n} \log(1 + 2\,\Re(\lambda_k) + |\lambda_k|^2)\right) \tag{11}$$

with the conventions $\log(0) := -\infty$ and $\exp(-\infty) := 0$. Furthermore,

$$\alpha - \frac{\alpha^2/2}{1+\min(0,\alpha)} \le \log(1+\alpha) \le \alpha \quad \text{for } -1 < \alpha \in \mathbb{R} \tag{12}$$

with equalities if and only if $\alpha = 0$.

PROOF OF (11) AND (12). Using $\det(I+E) = \prod_{k=1}^{n}(1+\lambda_k) = \exp\left(\sum_{k=1}^{n} \log(1+\lambda_k)\right)$ we obtain

$$|\det(I+E)| = \left|\exp\left(\sum_{k=1}^{n} \log(1+\lambda_k)\right)\right| = \exp\left(\sum_{k=1}^{n} \Re(\log(1+\lambda_k))\right) = \exp\left(\sum_{k=1}^{n} \log(|1+\lambda_k|)\right),$$

and (11) follows. To prove (12) we use $-\log(1-\beta) = \beta + \sum_{k=2}^{\infty} \dfrac{\beta^k}{k} \le \beta + \dfrac{\beta^2/2}{1-\beta}$ for $\beta \in [0,1)$ implying

$$\alpha - \frac{\alpha^2/2}{1+\alpha} \le \log(1+\alpha) \quad \text{for } \alpha \in (-1, 0].$$

The function $f(x) := x - x^2/2 - \log(1+x)$ with $f'(x) = -x^2/(1+x)$ is strictly decreasing for positive real x and satisfies $f(0) = 0$. That implies the lower bound in (12), and the upper bound is trivial. □

PROOF OF THEOREM 1.1. The Schur triangular form [8, Theorem 2.3.1] $E = UTU^H$ with unitary U and triangular T with λ_k on the diagonal implies $\sum_{k=1}^{n} |\lambda_k|^2 = \sum_{k=1}^{n} |T_{kk}|^2 \le \mathrm{tr}(T^H T) = \mathrm{tr}(E^H E) = \epsilon^2$, and the upper bound (3) follows by (11), (12), and

$$\begin{aligned}
\log|\det(I+E)| &\le \tfrac{1}{2}\sum_{k=1}^{n} 2\,\Re(\lambda_k) + |\lambda_k|^2 \\
&= \Re(\mathrm{tr}(E)) + \tfrac{1}{2}\sum_{k=1}^{n} |\lambda_k|^2 \\
&\le \Re(\mathrm{tr}(E)) + \tfrac{1}{2}\epsilon^2.
\end{aligned}$$

For the lower bound, $1 + 2\,\Re(\lambda_k) + |\lambda_k|^2 \ge (1+\Re(\lambda_k))^2$, (11) and (12) imply

$$\begin{aligned}
\log|\det(I+E)| &\ge \sum_{k=1}^{n} \log(1+\Re(\lambda_k)) \\
&\ge \Re(\mathrm{tr}(E)) - \sum_{k=1}^{n} \frac{(\Re(\lambda_k))^2/2}{1+\mu_k} \\
&\ge \Re(\mathrm{tr}(E)) - \tfrac{1}{2}\left(1+\min_k \mu_k\right)^{-1} \sum_{k=1}^{n} (\Re(\lambda_k))^2 \\
&\ge \Re(\mathrm{tr}(E)) - \frac{\epsilon^2/2}{1+\min_k \mu_k}.
\end{aligned}$$

The lower bounds in (5) and (6) follow by $\min_k \mu_k \ge -\rho(E) \ge -\epsilon$. The denominator in the lower bound of (6) cannot be replaced by 2 as shown by $E := \begin{pmatrix} 0 & \alpha \\ \alpha & 0 \end{pmatrix}$ with $|\alpha| < 1/\sqrt{2}$. □

Proof of (7). By (12) Hadamard's bound satisfies

$$\begin{aligned}\log|\det(I+E)| &\le \log\left(\textstyle\prod_{k=1}^n \|(I+E)_{k*}\|_2\right) = \frac{1}{2}\sum_{k=1}^n \log\left(1+2\,\Re(E_{kk})+\sum_{i=1}^n |E_{ki}|^2\right) \\ &\le \textstyle\sum_{k=1}^n \left(\Re(E_{kk})+\frac{1}{2}\displaystyle\sum_{i=1}^n |E_{ki}^2|\right) = \Re(\mathrm{tr}(E))+\frac{1}{2}\|E\|_F^2\end{aligned} \tag{13}$$

with the interpretation $\log(0)=-\infty$. It also follows that Hadamard's bound and the upper bound in (7) coincide if and only if $\Re(E_{kk})+\frac{1}{2}\sum_{i=1}^n |E_{ki}^2| = 0$ for all k. An example is $E=\begin{pmatrix} -\alpha & \sqrt{2\alpha-\alpha^2} \\ \sqrt{2\alpha-\alpha^2} & -\alpha \end{pmatrix}$ for $0\le\alpha\le 2$. □

Proof of Theorem 1.2. Let λ_k denote the eigenvalues of E. Then $|\lambda_k|\le\rho(E)<1$ implies

$$\det(I+E)=\exp\left(\sum_{k=1}^n \log(1+\lambda_k)\right)=\exp\left(\mathrm{tr}(E)+\sum_{k=1}^n \frac{\lambda_k^2}{2}\left(\sum_{j=0}^\infty \frac{(-1)^{j+1}2\lambda_k^j}{j+2}\right)\right) =: \exp(\mathrm{tr}(E)+\Phi).$$

Furthermore,

$$|\Phi| \le \sum_{k=1}^n \frac{|\lambda_k|^2}{2}\left(\sum_{j=0}^\infty |\lambda_k|^j\right) \le \frac{\epsilon^2}{2(1-\rho(E))} =: \Psi < 1$$

by $\rho(E)<1-\epsilon^2/2$. Hence, [10, 4.5.16] $|e^z-1|\le e^{|z|}-1$ for $z\in\mathbb{C}$ and [10, 4.5.11] $e^x-1\le\frac{x}{1-x}$ for $x<1$ give

$$\left|\frac{\det(I+E)-\exp(\mathrm{tr}(E))}{\exp(\mathrm{tr}(E))}\right| = |\exp(\Phi)-1| \le \exp(|\Phi|)-1 \le \frac{|\Phi|}{1-|\Phi|} \le \frac{|\Psi|}{1-|\Psi|} = \frac{\epsilon^2}{2(1-\rho(E)-\epsilon^2/2)}.$$

This implies (8). To show (9) note that (13) implies

$$\prod_{k=1}^n \|(I+E)_{k*}\|_2 \le |\exp(\mathrm{tr}(E))|\exp(\epsilon^2/2),$$

so that $e^x\le 1+\frac{x}{1-x}$ for $x:=\epsilon^2/2<1$ finishes that part. To see (10), we use $|\lambda_k|\le\epsilon<1$ and

$$|\Phi| \le \sum_{k=1}^n |\lambda_k^2|\left(\sum_{j=0}^\infty \frac{|\lambda_k|^j}{j+2}\right) \le \sum_{k=1}^n |\lambda_k^2|\left(\frac{1}{2}+\frac{|\lambda_k|}{3}+\frac{|\lambda_k|^2}{4(1-|\lambda_k|)}\right) \le \epsilon^2\left(\frac{1}{2}+\frac{\epsilon}{3}+\frac{\epsilon^2}{4(1-\epsilon)}\right).$$

Surely $|\Phi|<1$ for $\epsilon<0.7$, so that $\frac{|\Phi|}{1-|\Phi|}\le\frac{\epsilon^2}{2(1-\epsilon)}$ is equivalent to $(2-2\epsilon+\epsilon^2)|\Phi|\le\epsilon^2$. Now

$$(2-2\epsilon+\epsilon^2)|\Phi| \le \frac{\epsilon^2}{12(1-\epsilon)}(12-16\epsilon+8\epsilon^2-\epsilon^4) = \epsilon^2\left(1-\frac{(4-8\epsilon+\epsilon^3)\epsilon}{12(1-\epsilon)}\right)$$

is less than ϵ^2 if $\epsilon^3-8\epsilon+4>0$, and (10) follows. The upper bound $\epsilon^2/2$ one might want is not true as shown by any negative 1×1-matrix E satisfying the assumptions. □

Acknowledgement. The author wishes to thank Prashant Batra, Richard Brent, Florian Bünger and Marko Lange for fruitful comments and discussions. In particular my thanks to Richard Brent for pointing to the factor 1/2 in the lower bound in (12), and to Prashant Batra for pointing to the papers by Ostrowski [14] and Hans Schneider [16]. Moreover, I would like to thank in particular the referee for constructive and very valuable comments.

REFERENCES

[1] T. Bhatia and T. Jain, *Higher order derivatives and perturbation bounds for determinants*, LAA, 431 (2009), pp. 2102–2108.

[2] R. Brent, J. Osborn, and W. Smith, *Bounds on determinants of perturbed diagonal matrices.* arXiv:1401.7084v7, 18pp., 2014.

[3] ———, *Note on best possible bounds for determinants of matrices close to the identity matrix*, LAA, 466 (2015), pp. 21–26.

[4] L. Elsner, *Bounds for determinants of perturbed M-matrices*, LAA, 257 (1997), pp. 283–288.

[5] I. Fredholm, *Sur une classe d'équations fonctionelles*, Acta Math., 27 (1903), pp. 365–390.

[6] S. Gershgorin, *Über die Abgrenzung der Eigenwerte einer Matrix*, Izv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk, 6 (1931), pp. 749–754.

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

[8] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, second edition, 2017.

[9] I. Ipsen and R. Rehman, *Perturbation bounds for determinants and characteristic polynomials*, SIAM J. Matrix Anal. Appl. (SIMAX), 30 (2008), pp. 762–776.

[10] Frank W. Olver and Daniel W. Lozier and Ronald F. Boisvert and Charles W. Clark, *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, NY, USA, 1st edition, 2010.

[11] A. Ostrowski, *Sur la détermination des bornes inférieures pour une classe des déterminants*, Bull. Sci. Math. (2) 61, (1937), pp. 19–32.

[12] ———, *Über die Determinanten mit überwiegender Hauptdiagonale*, Comment. Math. Helv., 10 (1937), pp. 69–96.

[13] ———, *Sur l'approximation du déterminant de Fredholm par les déterminants des systèmes d'equations linéares*, Ark. Math. Stockholm Ser. A, 26 (1938), pp. 1–15.

[14] ———, *Note on bounds for determinants with dominant principal diagonal*, Proc. Amer. Math. Soc, 3 (1952), pp. 26–30.

[15] G. Price, *Bounds for determinants with dominant principal diagonal*, Proc. Amer. Math. Soc., 2 (1951), pp. 497–502.

[16] H. Schneider, *An inequality for latent roots of a matrix applied to determinants with dominant main diagonal*, Journal London Math. Soc., 28 (1953), pp. 8–20.