JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematics and Mathematical Physics
Vol. 65B, No. 1, January–March 1961

Bound for the P-Condition Number of Matrices With Positive Roots

Philip J. Davis, Emilie V. Haynsworth, and Marvin Marcus

(September 27, 1960)

Let a matrix A have positive roots $0<\lambda_1\leq\lambda_2\leq \ldots \leq\lambda_n$. Upper and lower bounds for the P-condition number of A, $P=\lambda_n/\lambda_1$, are given in terms of det A and one other symmetric function of the roots.

1. Introduction

Suppose $A=(a_{ij})$ is a nonsingular $n\times n$ matrix with roots, $\lambda_i(i=1, \ldots, n)$, ordered so that

$$0<|\lambda_1|\leq|\lambda_2|\leq\ldots\leq|\lambda_n|.$$

For a wide class of matrices A, the ratio $|\lambda_n|/|\lambda_1|$ gives a rough measure of the probable accuracy of the computation of the inverse of A, or the solution of the system of equations for which A is the matrix of coefficients.

This measure was evaluated in some detail by von Neumann and Goldstine [5][1] and has been called by J. Todd [2, 3, 4] the P- condition number of the matrix, i.e.,

$$P=\frac{|\lambda_n|}{|\lambda_1|}.$$

In general the accuracy of the results is in proportion to the reciprocal of P.

In this paper we shall show that, if the roots of A are all positive with

$$0<\lambda_1\leq\lambda_2\leq\ldots\leq\lambda_n \tag{1}$$

and the determinant of A, det A, is known, together with one other symmetric function of the roots, upper and lower bounds for P can be given in terms of these two known quantities.

More precisely, if the characteristic polynomial for the matrix A is

$$p(x)=x^n-c_1x^{n-1}+c_2x^{n-2}-\ldots+(-1)^n c_n,\ c_i\geq 0 \qquad (i=1, \ldots, n) \tag{2}$$

and we know c_k and $c_n=(\det A)$, there exist positive upper and lower bounds for the ratio

$$P=\frac{\lambda_n}{\lambda_1}, \tag{3}$$

in terms of c_k and c_n.

To simplify the notation in the statement and proof of the theorem we use the following device.

[1] Figures in brackets indicate the literature references at the end of this paper.

Let

$$c_k=c_k(\lambda_1, \ldots, \lambda_n)=\binom{n}{k} s_k(\lambda_1 \ldots, \lambda_n)=\binom{n}{k}s_k; \tag{4}$$

and divide each root of $p(x)$ by $(s_k)^{1/k}$. (This is equivalent to dividing each element of the matrix by $(s_k)^{1/k}$.) We will call the new matrix "normalized with respect to k" or simply the normalized matrix. Since the condition number, P, is the ratio of two roots, P will not be affected by such a transformation. Let D_k be the determinant of the normalized matrix. Then

$$D_k=\frac{s_n}{(s_k)^{n/k}}=\frac{\binom{n}{k}^{n/k}c_n}{c_k^{n/k}}. \tag{5}$$

From Hardy-Littlewood-Pólya [1] the numbers s_k as defined in (4) satisfy the inequalities

$$s_1\geq s_2^{1/2}\geq s_3^{1/3}\geq\ldots\geq s_n^{1/n} \tag{6}$$

(where equality holds only if all the λ_i are equal) hence, by (5), $D_k\leq 1$.

2. Statement and Proof of Theorem

THEOREM: *If we have a matrix $A=(a_{ij})$ with characteristic polynomial (2) and if the constant term and the kth coefficient are known, the following bounds hold for P,*

$$\frac{1}{D_k^{1/(n-1)}}\leq P\leq\frac{1+\sqrt{1-D_k^{n-1}}}{1-\sqrt{1-D_k^{n-1}}}, \qquad (k=1, \ldots, n-1) \tag{7}$$

where D is defined by (4) and (5).

If $k=1$ (i.e., the trace of the matrix is given) we can improve the upper bound:

$$P\leq\frac{1+\sqrt{1-D_1}}{1-\sqrt{1-D_1}}; \qquad D_1=\frac{n^nc_n}{c_1^n}. \tag{8}$$

PROOF: *Part 1—Upper bound.* We prove (8) first, as the method used for this can be applied in the proof of (7).

Since $P=\lambda_n/\lambda_1$, eq (8) is equivalent to

$$D_1\leq\frac{4\lambda_1\lambda_n}{(\lambda_1+\lambda_n)^2} \tag{9}$$

so from (5), we must show that

$$\frac{(\lambda_1+\lambda_n)^2}{4\lambda_1\lambda_n} \leq \left(\frac{s_1}{n\sqrt{s_n}}\right)^n. \tag{10}$$

But the right side of (10) is $(A_n/G_n)^n$ where A_n is the arithmetic mean, and G_n the geometric mean, of $\lambda_1, \ldots, \lambda_n$.

Thus, we can prove (10) by induction if we prove the following:

LEMMA:

$$\left(\frac{A_n}{G_n}\right)^n \leq \left(\frac{A_{n+1}}{G_{n+1}}\right)^{n+1}, \tag{11}$$

where A_{n+1} and G_{n+1} are the arithmetic and geometric means of $\lambda_1, \ldots, \lambda_n, \lambda_{n+1}$. (The λ_i are not assumed to be ordered with respect to size for this lemma.)

PROOF OF LEMMA: Simplifying (11) we see we must prove

$$\lambda_{n+1}A_n^n \leq A_{n+1}^{n+1}. \tag{12}$$

But (12) follows from the arithmetic-geometric mean inequality (cf., Hardy, Littlewood, Pólya [1]),

$$a_1a_2 \ldots a_{n+1} \leq \left(\frac{a_1+a_2+ \ldots +a_{n+1}}{n+1}\right)^{n+1},$$

if we let $a_i=A_n$, $(i=1, \ldots, n)$ and $a_{n+1}=\lambda_{n+1}$.

Thus (11) holds and (10) follows immediately using finite induction on n.

To prove (7) we note that, by the inequalities (6),

$$D_k^{n-1} \leq D_{n-1}^{n-1} = \frac{s_n^{n-1}}{s_{n-1}^n},$$

so it suffices to show that

$$\frac{(\lambda_1+\lambda_n)^2}{4\lambda_1\lambda_n} \leq \frac{1}{D_{n-1}^{n-1}} = \left(\frac{s_{n-1}}{s_n}\right)^n \cdot s_n. \tag{13}$$

Now let $\mu_1=\lambda_n^{-1}, \ldots, \mu_n=\lambda_1^{-1}$ and we note that $0<\mu_1\leq \ldots \leq \mu_n$ and (13) becomes

$$\frac{(\mu_1+\mu_n)^2}{4\mu_1\mu_n} \leq \left(\frac{s_1(\mu_1 \ldots \mu_n)}{s_n^{1/n}(\mu_1 \ldots \mu_n)}\right)^n. \tag{14}$$

But (14) is the same as (10) with the μ's substituted for the λ's, hence (13) holds and the general upper bound (7) is true for all values of k.

PROOF: *Part 2—Lower Bound.* From the inequalities (6),

$$D_k = \frac{s_n}{s_k^{n/k}} \geq \frac{s_n}{s_1^n},$$

where, by the definition (4)

$$s_1 \leq \lambda_n, s_n \geq \lambda_1^{n-1}\lambda_n.$$

Hence

$$\frac{1}{D_k} \leq \frac{s_1^n}{s_n} \leq \frac{\lambda_n^n}{\lambda_1^{n-1}\lambda_n} = \left(\frac{\lambda_n}{\lambda_1}\right)^{n-1}$$

which proves the lower bound in (7).

3. Some Remarks

In reference [6] T. Kato gives an upper bound for the P-condition number which in our notation is

$$P < \frac{4}{D_1}.$$

This is larger than the bound given in (8).

We derive several consequences from (7), (8). Suppose that we have a matrix with positive roots, which is normalized so that $c_k=\binom{n}{k}$, i.e., $s_k=1$. Then $D_k=\det A$, and we can write

$$\frac{1}{(\det A)^{1/(n-1)}} \leq P \leq \frac{1+\sqrt{1-(\det A)^{n-1}}}{1-\sqrt{1-(\det A)^{n-1}}}. \tag{15}$$

This implies that

$$\det A \to 0 \quad \text{if and only if } P\to\infty,$$

$$\det A \to 1 \quad \text{if and only if } P\to 1.$$

Hence, the det A behaves essentially as the reciprocal of the P-condition number and may be said to constitute a reasonable condition number of its own for such matrices. This lends substance to the popular feeling that for a properly defined class of normalized matrices the smallness of its determinant is accompanied by difficulty in inversion.

A second observation relates to least square approximations. Express the general problem in the language of inner product spaces. We are required to solve $\|y-\Sigma_{i=1}^n a_ix_i\|=$minimum, where y is given and x_i are independent elements. The normal equations have matrix $((x_i,x_j))$ where (p,q) designates the inner product. We can assume x_i normalized: $(x_i,x_i)=\|x_i\|^2=1$. The Gram matrix $((x_i,x_j))$ is positive definite symmetric and hence falls within the scope of our inequality with $k=1$, $c_1=n$. Hence if G is the Gram determinant $|(x_i,x_j)|$ we have

$$\frac{1}{G^{1/(n-1)}} \leq P \leq \frac{1+\sqrt{1-G}}{1-\sqrt{1-G}}. \tag{16}$$

The quantity G, which acts as a "measure" of linear independence of $x_1, \ldots, x_n$, therefore also serves as a condition number for the normal equations.

4. References

[1] G. H. Hardy, J. E. Littlewood, and G. Polya, Inequalities (Camb. Univ. Press, Cambridge, England 1952).

[2] J. Todd, The Condition of a Certain Matrix, Proc. Cambridge Phil. Soc. **46**, pp. 116–118 (1949).

[3] J. Todd, The condition of certain matrices I, Quart. J. Mech. Appl. Math. **2**, pp. 469–472 (1949).

[4] J. Todd, The condition of certain matrices II, Arch. Math **5**, pp. 249–257 (1954).

[5] J. von Neumann and H. Goldstine, Numerical inverting of matrices of high order, Bull. Am. Math. Soc. **53**, pp. 1021–1099 (1947).

[6] T. Kato, Estimation of Iterated Matrices, with application to the von Neumann condition, Numerische Math. **2**, pp. 22–29 (1960).

(Paper 65B1–42)