UNCLASSIFIED

CENTRAL RESEARCH LIBRARY
DOCUMENT COLLECTION

MARTIN MARIETTA ENERGY SYSYEMS LIBRARIES
3 4456 0323413 1

ORNL 1756
Physics Cy. 4

ON NORMS OF VECTORS AND MATRICES

A. S. Householder

CENTRAL RESEARCH LIBRARY
DOCUMENT COLLECTION

LIBRARY LOAN COPY

DO NOT TRANSFER TO ANOTHER PERSON

If you wish someone else to see this document, send in name with document and the library will arrange a loan.

OAK RIDGE NATIONAL LABORATORY
OPERATED BY
CARBIDE AND CARBON CHEMICALS COMPANY
A DIVISION OF UNION CARBIDE AND CARBON CORPORATION
UCC
POST OFFICE BOX P
OAK RIDGE, TENNESSEE

UNCLASSIFIED

ORNL 1756

Copy No. 4

Contract No. W-7405-eng-26

MATHEMATICS PANEL

ON NORMS OF VECTORS AND MATRICES

A. S. Householder

DATE ISSUED

AUG 26 1954

OAK RIDGE NATIONAL LABORATORY
Operated by
CARBIDE AND CARBON CHEMICALS COMPANY
A Division of Union Carbide and Carbon Corporation
Post Office Box P
Oak Ridge, Tennessee

MARTIN MARIETTA ENERGY SYSYEMS LIBRARIES

3 4456 0323413 1

INTERNAL DISTRIBUTION

1. C. E. Center
2. Biology Library
3. Health Physics Library
4-5. Central Research Library
6. Reactor Experimental Engineering Library
7-11. Laboratory Records Department
12. Laboratory Records, ORNL R.C.
13. C. E. Larson
14. L. B. Emlet (K-25)
15. J. P. Murray (Y-12)
16. A. M. Weinberg
17. E. H. Taylor
18. E. D. Shipley
19-143. A. S. Householder
144. F. C. VonderLage
145. C. P. Keim
146. W. H. Jordan
147. J. A. Swartout
148. S. C. Lind
149. F. L. Culler
150. R. S. Livingston
151. A. H. Snell
152. A. Hollaender
153. M. T. Kelley
154. G. H. Clewett
155. J. H. Frye, Jr.
156. K. Z. Morgan
157. T. A. Lincoln
158. C. S. Harrill
159. D. W. Cardwell
160. C. E. Winters
161. E. M. King
162. Lewis Nelson
163. D. D. Cowen
164. J. A. Lane
165. R. A. Charpie
166. S. J. Cromer
167. M. J. Skinner

EXTERNAL DISTRIBUTION

168. R. F. Bacher, California Institute of Technology
169. Division of Research and Medicine, AEC, ORO
170-422. Given distribution as shown in TID-4500 under Physics category

DISTRIBUTION PAGE TO BE REMOVED IF REPORT IS GIVEN PUBLIC DISTRIBUTION

ON NORMS OF VECTORS AND MATRICES

By a norm of a vector x will be meant any real valued function $\|x\|$ of x possessing the following three properties:

(I_v) $\|0\| = 0$, and for $x \neq 0$, $\|x\| > 0$;

(II_v) $\|\gamma x\| = |\gamma| \cdot \|x\|$ for any scalar γ;

(III_v) $\|x + y\| \leqslant \|x\| + \|y\|$.

By a norm of a square matrix A will be meant any real-valued function $\|A\|$ of A possessing the following four properties:

(I_m) $\|0\| = 0$, and for $A \neq 0$, $\|A\| > 0$;

(II_m) $\|\gamma A\| = |\gamma| \cdot \|A\|$ for any scalar γ;

(III_m) $\|A + B\| \leqslant \|A\| + \|B\|$;

(IV_m) $\|A B\| \leqslant \|A\| \cdot \|B\|$.

Following (2), a given matrix norm will be said to be consistent with a given vector norm provided for every A and every x (square matrices of order n and n-dimensional column vectors will be presupposed throughout) it is true that

$$(V) \qquad \|A x\| \leqslant \|A\| \cdot \|x\| .$$

If it is further true that for every A there exists an $x \neq 0$ such that

$$\text{(VI)} \qquad \|A\,x\| = \|A\| \cdot \|x\| ,$$

then the matrix norm is said to be subordinate to the vector norm. Faddeeva shows that given a vector norm a matrix norm can be defined by the condition

$$\|A\| = \max_{x \neq 0} \|A\,x\| \,/\, \|x\| .$$

We mention this in passing but shall not utilize the theorem here.

For the vector norm defined by

$$\|x\|^2 = x^T x ,$$

it is well known that

$$N(A) = \text{trace}\,(A^T A)$$

is a consistent but not a subordinate matrix norm. The subordinate matrix norm is equal to the square root of the largest proper value of $A^T A$. Faddeeva, and, previously, Bowker (1), discuss this norm, and also two others more readily computed. The purpose of the present note is to call attention to two families of norms which include these last two, as special cases. These classes were suggested by a recent note by de la Garza (3).

Before defining these families it may be worth while to summarize some of the more important properties of norms in general. In this it will be supposed merely that the matrix norm is consistent with the vector norm, but not necessarily that it is subordinate, unless the fact is mentioned explicitly. It is clear that among all matrix norms consistent with a given vector norm, the subordinate norm may be said to be optimal. The properties listed here are of some importance in numerical analysis (see 4).

If λ is any proper value of A, then

$$(1) \qquad |\lambda| \leqslant \|A\| .$$

For if x is an associated proper vector,

$$A x = \lambda x ,$$

then

$$\|A x\| = |\lambda| \cdot \|x\| \leqslant \|A\| \cdot \|x\| .$$

Since $x \neq 0$, $\|x\| \neq 0$, and the theorem follows.

If $\|H\| < 1$, then the series on the right of

$$(2) \qquad (I - H)^{-1} = I + H + H^2 + \ldots$$

converges to the left number. In fact, for any n,

$$(I - H)(I + H + \dots + H^n) = I - H^{n+1},$$

and since $\|H\| < 1$, therefore

$$\|H^{n+1}\| \leqslant \|H\|^{n+1}$$

approaches zero, whence H^{n+1} approaches the null matrix.

If $\|H\| < 1$, then

$$\|(I - H)^{-1}\| \leqslant \|I\| + \|H\| / (1 - \|H\|). \tag{3}$$

If the matrix norm is subordinate to the vector norm, then

$$\|I\| = 1. \tag{4}$$

In this event, if $\|H\| < 1$, then

$$\|(I - H)^{-1}\| \leqslant 1 / (1 - \|H\|). \tag{5}$$

To prove (4) we have only to note that there exists an $x \neq 0$ for which

$$\|I x\| = \|I\| \cdot \|x\|.$$

Then (5) follows almost immediately.

If C_0 is an approximate inverse of A, then

$$C_1 = C_0 (2 I - A C_0) = (2 I - C_0 A) C_0$$

is a better approximation. For convergence of the iteration it is sufficient that

$$\| H \| < 1 , \quad H = I - A C_0 ,$$

or that

$$\| K \| < 1 , \quad K = I - C_0 A .$$

When the first condition holds, then

$$\| A^{-1} \| \leqslant \| C_0 \| / (1 - \| H \|) , \tag{6}$$

and a similar inequality holds in the other case. This relation follows from the fact that

$$A^{-1} = C_0 (I - H)^{-1} ,$$

readily verified from the definition of H. From the definition of C, we have

$$C_1 = C_0 (I + H) = A^{-1} (I - H^2) ,$$

and, in general,

$$C_p = A^{-1} (I - H^{2^p}) .$$

Hence

$$A^{-1} - C_p = A^{-1} H^{2^p} ,$$

$$\| A^{-1} - C_p \| \leqslant \| A^{-1} \| \cdot \| H \|^{2^p} .$$

This demonstrates the convergence, and gives a measure of the rapidity of convergence.

A large class of iterative methods for solving the system

$$A x = y$$

is of the following character. Write

$$A = A_1 + A_2$$

where A_1 is nonsingular and easily inverted. Then the iteration

$$A_1 x_{p+1} = y - A_2 x_p ,$$

$$x_{p+1} = A_1^{-1} y - A_1^{-1} A_2 x_p$$

may converge to a solution. If

$$H = - A_1^{-1} A_2 ,$$

then a sufficient condition for convergence is that $\|H\| < 1$. Let

$$s_p = x - x_p$$

represent the deviation of any iterate from the true solution. Then

$$s_p = H s_{p-1} = H^p s_0 . \tag{7}$$

Hence

$$\|s_p\| \leqslant \|H\|^p \|s_0\|$$

which demonstrates the convergence.

Since

$$s_{p+1} = H s_p ,$$

$$s_p - s_{p+1} = (I - H) s_p ,$$

or

$$\begin{aligned} s_p &= (I - H)^{-1} (s_p - s_{p+1}) \\ &= (I - H)^{-1} H (s_{p-1} - s_p) , \\ s_p &= (I - H)^{-1} H (x_p - x_{p-1}) . \end{aligned}$$

Consequently

$$(8) \qquad \|s_p\| \leqslant \|H\| \cdot \|x_p - x_{p-1}\| / (1 - \|H\|) .$$

From this formula one can estimate the residual error by comparing the current and previous iterates.

Suppose one has a solution x of the system

$$A x = y$$

and wishes to determine the magnitude of the effect on x brought about by small changes in A and y. Thus one has (cf. 5),

$$(A - E)(x + d) = y + e ,$$

or

$$(A - E) d = e + E x ,$$

or

$$(I - H) d = A^{-1} (e + E x),$$

$$H = A^{-1} E .$$

Hence, finally

$$d = (I - H)^{-1} A^{-1} (e + E x).$$

Assuming, as usual, that $\| H \| < 1$, one has therefore

$$\| d \| \leqslant \| A^{-1} \| \; \| e + E x \| / (1 - \| H \|). \tag{9}$$

Turn, now, to the family of norms to be defined. Let g be any vector all of whose elements are positive (non-null):

$$g = (\gamma_i) \quad , \quad \gamma_i > 0 . \tag{10}$$

Such a vector will be called admissible. Let $| x |$ represent the vector whose elements are $| \xi_i |$, and let $| A |$ represent the matrix whose elements are $| \alpha_{ij} |$ (de la Garza calls this the abmatrix of A). Let a_i represent the columns of A. Then the g-norm of x is defined as

$$\| x \|_g = g^T | x | = \Sigma \gamma_i | \xi_i | , \tag{11}$$

and the g-norm of A is defined as

$$\| A \|_g = \max_j \| a_j \|_g / \gamma_i = \max_j \gamma_j^{-1} \sum_i \gamma_i | \alpha_{ij} | . \tag{12}$$

First it will be shown that these are norms. Since every $\gamma_i > 0$, properties I_v and I_m are obvious, while II_v and II_m follow from the linearity and the use of absolute values. To verify III_v one notes merely that

$$\sum \gamma_i \left| \xi_i + \eta_i \right| \leqslant \sum \gamma_i \left| \xi_i \right| + \sum \gamma_i \left| \eta_i \right| ,$$

and then applies the definition. Now III_v implies

$$\left\| a_i + b_i \right\|_g \leqslant \left\| a_i \right\|_g + \left\| b_i \right\|_g ,$$

for any i. Hence for any i

$$\begin{aligned} \gamma_i^{-1} \left\| a_i + b_i \right\|_g &\leqslant \gamma_i^{-1} \left\| a_i \right\|_g + \gamma_i^{-1} \left\| b_i \right\|_g \\ &\leqslant \max_i \gamma_i^{-1} \left\| a_i \right\|_g + \max_i \gamma_i^{-1} \left\| b_i \right\|_g \\ &= \left\| A \right\|_g + \left\| B \right\|_g . \end{aligned}$$

Since this holds for any i, it holds when the left member is maximized, which verifies III_m.

To prove IV_m, note that

$$\left\| A B \right\|_g = \max_j g^T \left| A B_j \right| / \gamma_j .$$

But

$$g^T \left| A\, b_j \right| = \sum_i \gamma_i \left| \sum_k \alpha_{ik}\, \beta_{kj} \right|$$

$$\leqslant \sum_i \gamma_i \sum_k \left| \alpha_{ik} \right| \left| \beta_{kj} \right|$$

$$= \sum_k g^T \left| a_k \right| \left| \beta_{kj} \right|$$

$$= \sum_k \gamma_k^{-1}\, g^T \left| a_k \right| \cdot \gamma_k \left| \beta_{kj} \right|$$

$$\leqslant \left\| A \right\|_g \cdot g^T \left| b_j \right| .$$

This holds for any j. Hence

$$\gamma_j^{-1}\, g^T \left| A\, b_j \right| \leqslant \left\| A \right\|_g \cdot \gamma_j^{-1}\, g^T \left| b_j \right|$$

$$\leqslant \left\| A \right\|_g \cdot \left\| B \right\|_g .$$

One has now only to maximize the left member with respect to j.

To show that the norms are consistent form

$$\left\| A\, x \right\|_g = g^T \left| A\, x \right| = \sum_i \gamma_i \left| \sum_j \alpha_{ij}\, \xi_j \right|$$

$$\leqslant \sum \sum \gamma_i \left| \alpha_{ij} \right| \cdot \left| \xi_j \right|$$

$$= \sum g^T \, |a_j| \, |\xi_j|$$

$$= \sum \gamma_j^{-1} g^T \, |a_j| \cdot \gamma_j \, |\xi_j|$$

$$\leqslant \|A\|_g \; g^T \, |x|$$

$$= \|A\|_g \; \|x\|_g \; .$$

Finally to show that the matrix norm is subordinate, let j be such that

$$\|A\|_g = \gamma_j^{-1} \; g^T \, |a_j| \; .$$

If $x = e_j$, the jth unit vector, then

$$A \, e_j = a_j \; , \quad \text{whence}$$

$$\|A \, e_j\|_g = \|a_j\|_g = g^T |a_j| = \gamma_j^{-1} g^T \, |a_j| \cdot \gamma_j$$

$$= \|A\|_g \, g^T \, |e_j| = \|A\|_g \cdot \|e_j\|_g \; .$$

Thus for $x = e_j$, VI is satisfied.

One of the norms considered by Bowker and Faddeeva is the particular norm obtained on taking

$$g = e = \sum e_i \; .$$

Clearly, however, in all the applications listed above it is advantageous to select a vector g for which the norm of the matrix of interest (designated by H in the applications) is as small as possible. An algorithm is immediately available, however, by which, from a $g = g_0$, one can construct a g_1 such that, for a particular matrix H,

$$\|H\|_{g_1} \leqslant \|H\|_{g_0} \quad . \tag{13}$$

That is, for the matrix H, the g_1-norm is at worst no larger, and in general is smaller, than the g_0-norm. The algorithm is simply

$$g_1 = |H^T| \, g_0 \quad .$$

Thus one could, in principle, construct a decreasing sequence of norms by means of the iteration

$$g_{p+1} = |H^T| \, g_p \quad . \tag{14}$$

In verifying (13) it is convenient to let the relation

$$|x| \leqslant |y|$$

signify that

$$|\xi_i| \leqslant |\eta_i|$$

for every i, and let

$$|A| \leqslant |B|$$

signify that

$$|\alpha_{ij}| \leqslant |\beta_{ij}|$$

for every i and j. For brevity also let

$$\nu_0 = \|H\|_{g_0} .$$

Then ν_0 is the smallest scalar for which

$$g_0^T |H| \leqslant \nu_0 g_0^T .$$

Since all quantities are positive, this inequality implies that

$$g_0^T |H|^2 \leqslant \nu_0 g_0^T |H| ,$$

and this is to say that

$$g_1^T |H| \leqslant \nu_0 g_1^T .$$

Consequently, if ν_1 is the smallest scalar for which

$$g_1^T \, |H| \leqslant \nu_1 \, g_1^T ,$$

it follows that

$$\nu_1 \leqslant \nu_0 .$$

That in general the terms ν_0 , ν_1 , ν_2 , ... of the sequence defined by repetition of this process are not all equal is evident from the following considerations: The matrix $|H|$ has all its elements non-negative, and for such a matrix there is a real, positive proper value which is exceeded by no other proper value in magnitude. The corresponding proper vector has all its elements real and positive. If h is this proper vector, then the h-norm is optimal for H in the sense that

$$\|H\|_h \leqslant \|H\|_g$$

for all admissible vectors g . Moreover, as is well known, the sequence of vectors g_i approaches h as a limit.

Of the two families referred to at the outset, the other one is obtained quite obviously by placing the vector g on the right of the matrix. Thus, for any matrix A , let us now designate

its rows by a_i^T , and define now the norm $\|A\|_g$, by

$$(15) \qquad \|A\|_{g'} = \max_i \left| a_i^T \right|_g / \gamma_i ,$$

and the norm $\|x\|_{g'}$ by

$$(16) \qquad \|x\|_{g'} = \max_i \gamma_i^{-1} \left| \xi_i \right| .$$

That these are, in fact, norms, and that the matrix norm is consistent with and subordinate to the vector norm, can be proved very much as before.

As byproducts of these considerations, it may be noted that certain classical theorems can be proved and generalized almost immediately. We conclude by mentioning two.

The first states that a determinant with dominant principal diagonal is non-null (e.g., see 7). It was shown above that if $\|H\| < 1$, however the norm of H may be defined, then $I - H$ is nonsingular (since its inverse was obtained as the limit of a converging infinite series). If $\|H\| = \|H\|_e$, where $e = \Sigma e_i$, and H has only zeros in its diagonal, this proves the theorem for the special case of the determinant of a matrix $A = I - H$ with units along the diagonal. If D is an arbitrary diagonal matrix all cases are taken care of by considering $A = D(I - H)$ or $A = (I - H)D$. An obvious generalization results on considering $\|H\|_g$ for any g , and

the same holds for $\|H\|_{g'}$.

The other theorem is that of von Mises and Pollaczek-Geiringer (6), and, in fact, it was proved above. Their theorem, expressed in these terms, states merely that the iteration

$$x_{p+1} = y + H x_p ,$$

converges in case $\|H\|_e < 1$ or $\|H\|_{e'} < 1$. The generalization comes about, again, on replacing e by any other admissible vector g.

As a final note, mention may be made of a natural generalization of the norm $\|x\|$ defined by

$$\|x\|^2 = x^T x .$$

If G is any positive definite matrix the determinant

$$|A^T G A - \lambda G|$$

can be shown to have only non-negative real zeros. If λ_1 is the greatest of these the function

$$\|A\|_G = \sqrt{\lambda_1}$$

can be shown to a norm, and to be subordinate to the vector norm defined by

$$\|x\|_G = \sqrt{x^T G A} .$$

REFERENCES

1. Albert H. Bowker, On the norm of a matrix, An. Math. Stat. 18 (1947) 285-88.

2. V. N. Faddeeva, Computational methods of linear algebra (Russian). Moscow and Leningrad. 1950.

3. A. de la Garza, Error bounds on approximate solutions to systems of linear algebraic equations, MTAC 7 (1953) 81-84.

4. A. S. Householder, Principles of numerical analysis. New York. McGraw-Hill. 1953.

5. A. T. Lonseth, The propagation of error in linear problems. Tr. Am. Math. Soc. 62 (1947) 193-212.

6. R. von Mises and Hilda Pollaczek-Geiringer, Praktische Verfahren der Gleichungsauflosung. Z. Ang. Math. u. Mech. 9 (1929) 58-77, 152-64.

7. Olga Taussky, A recurring theorem on determinants, Am. Math. Monthly 56 (1949) 672-76.

mg