UNCLASSIFIED

CENTRAL RESEARCH LIBRARY
DOCUMENT COLLECTION

MARTIN MARIETTA ENERGY SYSTEMS LIBRARIES
3 4456 0349689 3

ORNL 1785
Physics

4

ON SOLVING LINEAR ALGEBRAIC SYSTEMS

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 1785

Copy No. 4

Contract No. W-7405-eng-26

MATHEMATICS PANEL

ON SOLVING LINEAR ALGEBRAIC SYSTEMS

A. S. Householder

DATE ISSUED

SEP 21 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 SYSTEMS LIBRARIES

3 4456 0349689 3

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
170-423. Given distribution as shown in TID-4500 under Physics category

DISTRIBUTION PAGE TO BE REMOVED IF REPORT IS GIVEN PUBLIC DISTRIBUTION

ON SOLVING LINEAR ALGEBRAIC SYSTEMS

Most methods of solving a linear algebraic system of equations

$$A\,x = y \tag{1}$$

fall into two broad classes, generally called iterative and direct. An iterative method, as the term is used here, defines an infinite sequence of vectors x_i which converge to the solution x, while a direct method leads in principle to the true solution after a finite number of steps. Ordinarily this is true in principle only, since errors generated in the course of the computation (due to rounding) will contaminate the result, so that the method may be expected to yield only an approximation that may or may not be sufficiently close in any given case.

An iterative method usually has two advantages over any direct method. These are, first, that each term in the sequence is obtainable in a relatively simple manner. For programming on an automatic computer, the simplicity of the program may be an important consideration. The second advantage is in part a consequence of the first and is this, that the generated error may be reducible to a minimum. This is true in particular in case each term x_p is obtained as an "improvement" over the preceding x_{p-1} by operating upon it alone. In assessing the precision that is achievable by such a method, it is sufficient to ask under what circumstances one is assured that the result of an iteration x_p^*, after having been rounded from x_p, will still be better than the x_{p-1}

previously at hand. How x_{p-1} had been obtained (presumably by rounding after a previous iteration) is irrelevant.

The two advantages just mentioned may, of course, be partially or completely offset by slow convergence. To handle the most recalcitrant cases, therefore, one might need to apply first a direct method for obtaining rapidly a reasonably close but possibly not adequate first approximation, and then to improve this by a few iterative cycles to eliminate the worst of the generated error. But if this is to be done one would like the iterative and direct methods employed to be in some manner related.

Supposing the matrix A to be arbitrary, the equations to be solved may be interpreted as requiring that the vector y be resolved along the vectors represented by the columns of A, the elements of x providing the coefficients. If x_0 is any approximate set of multipliers, and

$$r_0 = y - A x_0 \tag{2}$$

represents the discrepancy, a better approximation can be had by projecting r_0 upon any vector u_1 of the form $u_1 = A v_1$. One determines λ_1, that is, so that the vector

$$x_1 = x_0 + \lambda_1 v_1 \tag{3}$$

represents a new approximation whose residual

$$r_1 = r_0 - \lambda_1 u_1 = y - A x_1 \tag{4}$$

is orthogonal to u_1. Since r_0 is thus resolved into components r_1 and u_1, unless r_0 is itself orthogonal to u_1 the step leads to an actual improvement in the sense that $N(r_1) < N(r_0)$, where, by definition,

$$N(x) = (x^T x)^{1/2}$$

for any x (the notation throughout is that of Householder, 1953). But if

$$s_p = x - x_p ,$$

it is not necessarily the case that $N(s_1) < N(s_0)$.

The expression for λ_1 which satisfies the given requirements is simply

$$\lambda_1 = u_1^T r_0 / u_1^T u_1 . \tag{5}$$

If one proceeds, now, to select vectors v_2, v_3, v_4, ..., in sequence, obtaining therefrom the sequence x_2, x_3, x_4, ..., the sequence will converge in general, so long as the v's are not subject to rather obvious pathological restrictions.

If the v's are selected arbitrarily, except for not satisfying the pathological restrictions, then part of the information available from preceding steps is ignored. Thus, since r_1 was made orthogonal to u_1, it follows that the best selection of $u_2 = A v_2$ will necessarily lie in the space orthogonal to u_1. If such a selection is made, then r_2 will

be orthogonal to u_1 as well as u_2. Hence $u_3 = A\,v_3$ is best taken orthogonal to both u_1 and u_2. But, in following this process, if A is nonsingular, one will arrive eventually at $r_m = 0$ for some $m \leqslant n$, since r_n would be orthogonal to the n mutually orthogonal, and hence linearly independent, vectors $u_1, u_2, \ldots, u_n$.

To describe the process in quite general terms, let

$$P = (p_1, p_2, \ldots, p_n) \tag{6}$$

be a nonsingular matrix, and let

$$Q = A\,P = (q_1, q_2, \ldots, q_n)\,. \tag{7}$$

One can determine a unit upper triangular matrix M such that

$$P = V\,M\,, \quad Q = U\,M\,, \quad U^T U = D^2, \tag{8}$$

where D is a nonsingular diagonal matrix. An induction will both demonstrate the theorem and show how M, V and U are to be computed.

First one takes

$$p_1 = v_1\,, \quad q_1 = u_1\,.$$

Next, suppose

$$P_i = (p_1, p_2, \ldots, p_i)\,,$$

with Q_i , V_i , and U_i similarly defined, and suppose

$$P_i = V_i M_i , \qquad Q_i = U_i M_i , \qquad U_i^T U_i = D_i^2 .$$

The first two of these relations are equivalent. It is required now to form a bordering vector m_i so that

$$(Q_i , q_{i+1}) = (U_i , u_{i+1}) \begin{pmatrix} M_i & m_{i+1} \\ 0 & 1 \end{pmatrix}$$

and

$$(U_i , u_{i+1})^T (U_i , u_{i+1}) = \begin{pmatrix} D_i^2 & 0 \\ 0 & \delta_{i+1}^2 \end{pmatrix} .$$

This requires that

$$q_{i+1} = U_i m_{i+1} + u_{i+1} ,$$

$$U_i^T u_{i+1} = 0 ,$$

$$u_{i+1}^T u_{i+1} = \delta_{i+1}^2 .$$

The first two of these three relations give

$$U_i^T q_{i+1} = D_i^2 m_{i+1} , \quad m_{i+1} = D_i^{-2} U_i^T q_{i+1} .$$

Knowing m_{i+1} , one can now form u_{i+1} , v_{i+1} and δ^2_{i+1} .

As a byproduct of these operations, note that $U D^{-1}$ is an orthogonal matrix:

$$(U D^{-1})^T U D^{-1} = I .$$

Hence

$$U D^{-2} U^T = I .$$

Since

$$U = A V ,$$

therefore

$$I = U D^{-2} U^T = A V D^{-2} U^T ,$$

(9) $$A^{-1} = V D^{-2} U^T .$$

Successive steps of the process give approximations x_i that are progressively "better" in the sense described above. One has

(10) $$x_i = x_0 + V_i \ell_i , \quad r_i = r_0 - U_i \ell_i ,$$

$$\ell_i = D_i^{-2} U_i^T r_0 .$$

The improvement brought about in any step is measured by

$$(11) \qquad r_{i-1}^T r_{i-1} - r_i^T r_i = (u_i^T r_{i-1})^2 \delta_i^{-2} = (u_i^T r_0)^2 \delta_i^{-2} .$$

A convenient and natural choice for the matrix P is the identity. Then (8) becomes

$$(12) \qquad I = V M , \quad A = U M , \quad U^T U = D^2 ,$$

and the columns of U are the orthogonalized columns of A.

If there are only $m < n$ columns in A but these are linearly independent, the same process leads to an r_m which is orthogonal to all the columns of A, and hence to a solution of the least squares problem.

To return to the case $m = n$ (the more general case differs only trivially), one will find, on reaching r_n, that it does not vanish strictly but remains of a magnitude that may or may not be considered a serious deviation. The principal purpose of the present note, however, is to point out, first, that one has now readily available the framework for rapid improvement, and, second, to develop a formula that can be utilized to show at what stage any further "improvements" become submerged in the error that is generated in the course of computing them.

The method one uses to improve upon the x_n is simply this, that one utilizes the already available v_1, v_2, ..., and associated u_1, u_2, ..., in the simple iterative process. The fact that the vectors u_i are nearly orthogonal (but not strictly so because of the presence of generated errors), is advantageous for convergence. The fact that they are not

strictly orthogonal means merely that the exact solution cannot be achieved in n steps.

Now consider any $u = A\,v$, which may or may not be one of the u_i of the orthogonal set. Let x_p be any digital vector, regarded as an approximation to the true solution. How it may have been obtained is at present irrelevant. We wish to form

$$x_{p+1} = x_p + \lambda v ,$$

by selecting λ so that r_{p+1} and u are orthogonal. Actually we shall form

$$x^*_{p+1} = x_p + (\lambda^* v)^* .$$

Under what circumstances can we be assured that x^*_{p+1}, the vector we shall actually form, will be a better approximation than x_p ? The approximation will be better in fact, by the criterion being used here, when

$$N(r_{p+1}) < N(r_p) , \tag{13}$$

where

$$r_{p+1} = y - A\,x^*_{p+1} .$$

However, we have available neither r_p, nor r_{p+1}, but only

(14) $$r_p^* = y - (A\, x_p)^*$$

and

(15) $$r_{p+1}^* = y - (A\, x_{p+1}^*)^* .$$

Since, for any x and y,

$$N(y) \geqslant N(x) - N(x - y) ,$$

it follows that (13) is assured if

$$N(r_p^*) - N(r_p^* - r_p) > N(r_{p+1}^*) + N(r_{p+1}^* - r_{p+1}) .$$

Upper bounds for $N(r_p^* - r_p)$ and $N(r_{p+1}^* - r_{p+1})$ are the same and are fixed by the routine for carrying out the computations (14) and (15). If all elements of A and of both x's are digital, and each multiplication is rounded off, then

$$N(r_p^* - r_p) \leqslant n^{3/2}\, \epsilon ,$$

and likewise for the next residuals. In this event (13) will be assured by

(16) $$N(r_p^*) > N(r_{p+1}^*) + 2\, n^{3/2}\, \epsilon .$$

If complete products are accumulated and only the sums rounded the factor

$n^{3/2}$ can be replaced by $n^{1/2}$.

This is a test that can be applied to the two computed residuals after they have been computed, but they give no indication as to when one might expect the relation to hold.

To consider this we note first that

$$(17) \qquad (u^T r_p)^2/u^T u$$

is the amount that the squared norms of the residuals should differ if the strict arithmetic operation were carried out. Consider, then, the amount by which the generated error can affect this. This is simply

$$\left[y - A x_{p+1}\right] - \left[y - A x^*_{p+1}\right] = A(x^*_{p+1} - x_{p+1}) ,$$

where

$$x^*_{p+1} - x_{p+1} = (\lambda - \lambda^*)v + \lambda^* v - (\lambda^* v)^* .$$

Let

$$\overline{r}_{p+1} = y - A x^*_{p+1}$$

represent the true residual corresponding to the computed x^*_{p+1}. Then

$$(18) \qquad N(r_{p+1} - \overline{r}_{p+1}) \leqslant M(A) \left[|\lambda - \lambda^*| \, N(v) + n^{1/2} \epsilon \right] ,$$

where $M(A)$ is the square root of the largest proper value of $A^T A$. An upper bound for the proper value is provided by any norm of $A^T A$ [Householder, 1954]. <u>We are assured of a real improvement if v is such that the square root of the expression (17) exceeds twice the right member of (18).</u>

In (18) the bound for $|\lambda - \lambda^*|$ depends upon the vector v and upon the routine for calculating λ. Since

$$\lambda = u^T r_p / u^T u ,$$

we may assume that r_p is small compared to u and that $|\lambda| < 1$. For one possible routine the difference $\lambda - \lambda^*$ may be written

$$\lambda - \lambda^* = u^T r_p \left[\frac{u^T(u - u^*)}{(u^T u)(u^T u^*)} + \frac{(u^T - u^{*T})u^*}{(u^T u^*)(u^{*T} u^*)} + \frac{(u^{*T} u^*) - (u^{*T} u^*)^*}{(u^{*T} u^*)(u^{*T} u^*)^*} \right]$$

$$(19) \qquad + \frac{u^T(r_p - r_p^*) + (u^T - u^{*T}) r_p^* + u^{*T} r_p^* - (u^{*T} r_p^*)^*}{(u^{*T} u^*)^*}$$

$$+ \frac{(u^{*T} r_p^*)^*}{(u^{*T} u^*)^*} - \left[\frac{(u^{*T} r_p^*)^*}{(u^{*T} u^*)^*} \right]^* .$$

If $N(u^*)$ differs only negligibly from $N(u)$, by comparison with $N(u)$, this implies

$$(20) \quad |\lambda - \lambda^*| \leqslant \epsilon\, N^{-3}(u) \left\{ n \left[N(u) + N(r_p) \right] + n^{3/2}\, N(u) \left[N(u) + 3N(r_p) \right] + 2\, N^3(u) \right\}.$$

If $N(r_p)$ is small by comparison with $N(u)$, then approximately

$$(21) \qquad |\lambda - \lambda^*| \leqslant \epsilon \left[n\, N^{-2}(u) + n^{3/2}\, N^{-1}(u) + 2 \right] ,$$

and one of the two first terms may be expected to dominate.

In (20) and (21) it is presupposed that in forming scalar products of vectors, each product of elements is rounded. If complete products are accumulated and only the sums are rounded, then the factor n can be dropped and $n^{3/2}$ replaced by $n^{1/2}$.

When the matrix A is of rank n, and not too badly conditioned, the vectors r_p of the sequence and the scalars λ should have become small before the magnitude of $|\lambda - \lambda^*|$ becomes a factor to consider. Hence a further reduction in this quantity can be had at the expense of more programming. This is by computing $(2^\beta r_p)^*$, for suitable β [cf. Householder, 1952], and thence obtaining $(2^\beta \lambda)^*$.

Ultimately one is interested in estimates of the magnitude of s_p, rather than of r_p. But for any definition of norm, the norm of s_p can be bounded knowing the norm of r_p and the norm of A^{-1}, and this can, in turn, be estimated from any approximation to A^{-1} [cf. Householder, 1954].

The method developed above makes particular use of the column vectors of A. In fact, each x_i is obtained by projecting upon the space of the column vectors $a_1, a_2, \ldots, a_i$, so that the new residual

is orthogonal to this space. It should be noted that if $x_0 = 0$, and one takes $P = I$, $Q = A$ in equations (6) and (7), then only the first i elements of x_i are non-null, and, in fact, x_i represents the least squares solution of the equations from which the last $n - i$ terms are left out.

Consider now only the case when A is nonsingular, and let the adjoint system be written

$$A^T z = w \quad . \tag{22}$$

The results of the previous computations yield almost immediately the solution of the adjoint system (22). Otherwise put, if the system to be solved is expressed in the form (22), one may orthogonalize the rows of the matrix (thought of now as a transpose) instead of its columns.

In terms of the geometry of the system (22), one proceeds as follows: Given any initial approximation z_0 to the solution z, and a vector v_1 of multipliers, one projects the point z_0 orthogonally upon the plane

$$u_1^T z = v_1^T w \ , \qquad u_1 = A v_1 \ . \tag{23}$$

This plane contains the point z satisfying (22). The projection is expressible as

$$z_1 = z_0 + \mu_1 u_1 \ , \tag{24}$$

where μ_1 is chosen so that z_1 satisfies (23):

$$\mu_1 = v_1^T t_0 / u_1^T u_1 , \quad t_0 = w - A^T z_0 . \tag{25}$$

The new residual

$$t_1 = w - A^T z_1 = t_0 - \mu_1 A^T u_1 \tag{26}$$

is the projection of t_0, the initial residual, upon the plane (23), and is therefore not longer than t_0, and can be equal only if z_0 itself is in the plane (23).

One could select v_2 arbitrarily, project z_1 orthogonally upon the plane

$$u_2^T z = v_2^T w , \quad u_2 = A v_2 ,$$

and continue forming an infinite converging sequence of vectors z_p. But the plane (23) is known to contain both points z and z_1. Hence if u_2 is orthogonal to u_1, z_2 will also lie in this plane. Thus by using the orthogonal set u_i the iteration converges to the true solution in at most n steps.

Again, however, the z_n actually obtained will not be the strictly correct z, nor will the vectors u_i be strictly orthogonal. Nevertheless, using z_n as z_0', one can proceed with the now available, nearly orthogonal u_i, carrying out an iteration that should converge with some rapidity.

In this iteration the effectiveness of a given step is measured by

$$(27) \qquad N^2(t_p) - N^2(t_{p+1}) = u^T A A^T u \, (v^T t_p)^2 / (u^T u)^2 .$$

If z_p is any approximation (digital), and z^*_{p+1} the digital iterate obtained by use of the vector v, then

$$t_{p+1} - \overline{t_{p+1}} = (w - A^T z_{p+1}) - (w - A^T z^*_{p+1})$$

$$= A^T (z^*_{p+1} - z_{p+1}) ,$$

with

$$z_{p+1} - z^*_{p+1} = (\mu - \mu^*) u + u^*(u - \mu^*) + \mu^* u^* - (\mu^* u^*)^* .$$

Hence

$$(28) \qquad N(t_{p+1} - \overline{t_{p+1}}) \leqslant M(A) \Big[|\mu - \mu^*| \, N(u) + |\mu^*| \, N(u - u^*)$$

$$+ N\left\{\mu^* u^* - (\mu^* u^*)^*\right\} \Big] .$$

If each product is rounded, then

$$N\left\{\mu^* u^* - (\mu^* u^*)^*\right\} \leqslant n^{1/2} \epsilon ,$$

$$N(u - u^*) \leqslant n^{3/2} \epsilon .$$

Hence

$$(29)\quad N(t_{p+1} - \overline{t}_{p+1}) \leqslant M(A)\left[n^{1/2}\,\epsilon + |\mu^*|\; n^{3/2}\epsilon + |\mu - \mu^*|\; N(u)\right] .$$

The factor $\mu - \mu^*$ can be decomposed

$$\mu - \mu^* = \frac{(v^T t_p)\, u^T(u^* - u)}{(u^T u)(u^T u^*)} + \frac{(v^T t_p)(u^{*T} - u^T)u^*}{(u^T u^*)(u^{*T} u^*)}$$

$$+ \frac{(v^T t_p)\left[(u^{*T} u^*)^* - (u^{*T} u^*)\right]}{(u^{*T} u^*)(u^{*T} u^*)^*}$$

$$+ \frac{v^T(t_p - t_p^*) + v^T t_p^* - (v^T t_p^*)^*}{(u^{*T} u^*)^*}$$

$$+ \frac{(v^T t_p^*)^*}{(u^{*T} u^*)^*} - \left[\frac{(v^T t_p^*)^*}{(u^{*T} u^*)^*}\right]^* ,$$

which leads to

$$(30)\qquad |\mu - \mu^*| \leqslant \epsilon\, N^{-4}(u)\left\{ n\left[N(v)\; N(t_p) + N^2(u)\right] \right.$$

$$\left. + \; n^{3/2}\, N(u)\; N(v)\left[2N(t_p) + N(u)\right] + 2N^4(u)\right\} .$$

If $N(t_p)$ can be neglected, then

$$(31) \qquad |\mu - \mu^*| \leqslant \epsilon\, N^{-2}(u) \left[n + n^{3/2}\, N(v) + 2N^2(u) \right] .$$

When this is substituted into the right of (29) the result is to be compared with the square root of the right member of (27):

$$N(A^T u)\, |\, v^T t_p \,|\, N^{-2}(u) \quad .$$

Finally, when z_p is accepted as an adequate approximation, one can, if necessary, use the relation

$$(32) \qquad \|z - z_p\| \leqslant \|A^{-1}\| \cdot \|t_p\| \quad ,$$

valid for any consistent norms whatever [Householder, 1954] to estimate the magnitude of $z - z_p$.

At the possible risk of belaboring the obvious, a final word may be appended with respect to the criterion for terminating the iteration. This criterion relates to $N(r_p)$ or $N(t_p)$ as the case may be. On the other hand, the real objective is an x_p or a z_p for which $N(x - x_p)$ or $N(z - z_p)$ is as small as possible. It is known that the case

$$N(x - x_{p+1}) > N(x - x_p)$$

can arise, even though always

$$N(r_{p+1}) \leqslant N(r_p) \quad .$$

The fact, however, that this case can arise is entirely irrelevant, because its occurence can never be recognized in a given instance unless one has more information than is here presupposed. If such additional information were at hand, then another method of solution would probably be indicated. Lacking it, one can only estimate the error in z or in x by a relation such as (32), by using the t_p (or the r_p) which is in fact at hand.

Following the terminology of Hestenes and Stiefel (1952) the methods that yield x_n and z_n may be called "Methods of Orthogonal Directions", since they constitute natural extensions to more general matrices the methods for positive definite matrices which they call "Methods of Conjugate Directions". The point of view here proposed toward the subsequent iteration, and toward the assessment of the iterates, is different. In particular, deviations of the u's from strict orthogonality is of no great concern, nor does it require consideration in the analysis of errors.

These methods lead to two distinct generalizations of the method of conjugate gradients, one given by Hestenes and Stiefel and one not. The generalizations may be called Methods of Orthogonal Gradients, and presentation of both in the present geometric setting may be of interest.

For the first method, let

$$\varphi(x) = (y - A x)^T (y - A x) = r^T r . \tag{33}$$

This function is essentially positive and vanishes only for $x = A^{-1} y$. Its gradient is

$$\varphi_x = -2 A^T r \quad . \tag{34}$$

The method of steepest descent applied to this problem takes

$$x_{p+1} = x_p + \lambda_p v_p \quad ,$$

$$v_p = A^T r_p \quad .$$

This is a simple iterative method (non terminating). The method of orthogonal gradients would take for the matrix P of equation (6) and following

$$p_i = A^T r_{i-1} \quad .$$

This is the method described by Hestenes and Stiefel. The special advantage in this choice of the vectors p_i lies in the fact that the recursion for the orthogonalized u_i has the very simple form

$$u_i = q_i + \beta_{i-1} u_{i-1} \quad ,$$

all other terms dropping out and leaving only the single parameter β_{i-1} to be determined. That is to say, the matrix M in (8) has non-null elements along and just above the main diagonal, but nowhere else.

The other method of orthogonal gradients would make use of the adjoint. Let

$$\psi(z) = (w - A^T z)^T (w - A^T z) = t^T t ,$$

$$\psi_z = - 2 A t.$$

For the recursion which starts with (24), one takes

$$p_i = t_{i-1} .$$

The matrix M takes on the same simple form as in the other case.

REFERENCES

1. M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. J. R. N. B. S. 49(1952) 409-36.

2. A. S. Householder. Errors in iterative solutions of linear systems. Proceedings of the ACM, Toronto 1952.

3. ———————— Norms of vectors and matrices. ORNL Report 1756, 1954.