1963

Investigation of measures of ill-conditioning

Thomas D. Calton

INVESTIGATION

OF

MEASURES OF ILL-CONDITIONING

BY

THOMAS D. CALTON

A

THESIS

submitted to the faculty of the

SCHOOL OF MINES AND METALLURGY OF THE UNIVERSITY OF MISSOURI

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE, APPLIED MATHEMATICS

Rolla, Missouri

1963

Approved by

Ralph E. Lee (advisor) C A Johnson

Harold Q Fuller

TABLE OF CONTENTS

Page

ACKNOWLEDGMENT iii

LIST OF TABLES iv

LIST OF FIGURES v

I. INTRODUCTION. 1

II. REVIEW OF LITERATURE 3

III. DISCUSSION 12

IV. CONCLUSIONS 29

V. SUMMARY 32

VI. APPENDIX 33

BIBLIOGRAPHY 39

VITA 40

ACKNOWLEDGMENT

The author wishes to express his sincere appreciation to Professor Ralph E. Lee for his assistance and guidance toward the preparation of this investigation. Also, the author wishes to express gratitude to the Missouri School of Mines and Metallurgy Computer Center for the use of their equipment in the course of this thesis.

LIST OF TABLES

Page

I. MEASURES AND VALUES FOR VARIOUS SYSTEMS STUDIED. 26

LIST OF FIGURES

Page

1. Graph of variable coefficients and E-Multiples. 18

I. INTRODUCTION

The object of investigation in this study is that of ill-conditioning in the solution of a system of linear equations. Ill-conditioning arises when the solution is very sensitive to small changes in the coefficients of the unknowns.

A study is made in this paper of the various proposed measures of ill-conditioning for the purpose of finding the most practical method or measure for determining whether a system is ill-conditioned.

The problem of near-singular or ill-conditioned systems is of great importance in the solution of linear systems because of the extensive use made of them in practical situations. Solution of a system of linear equations with this property frequently finds use in many areas of Applied Science. In applied mathematics, systems of linear equations are used in solving such problems as method of least squares, solution of partial differential equations, ordinary differential equations and many others.

Although computation of such a system could be done by double precision, giving increased accuracy at each step, this does not eliminate the problem. The problem of obtaining accurate data may be more important than the actual computation. However, when a system is found to be ill-conditioned, a method of higher precision is often used to improve round-off errors which would invalidate the solution. Nevertheless, it is the identification or means of detecting such a system which needs to be considered before further analysis can be pursued. In small systems the detection of ill-conditioning is fairly obvious by observation; whereas for larger systems, it is hidden from observation in most cases. Thus, an indicative measure is needed to detect such a system.

It is the aim of the author in this study to find a suitable measure or method for detecting an ill-conditioned system of equations.

II. REVIEW OF LITERATURE

An examination of the literature available on the subject of simultaneous-linear equations reveals several proposed measures or tests to establish whether or not a system of equations is ill-conditioned.

The problem of ill-conditioning is an obstacle in the solution of simultaneous linear equations which can be most serious and result in a solution that has no meaning at all with respect to the particular situation it is to predict or describe. This was illustrated by Macon[1] with the following system of equations whose solution is obvious

$$\begin{aligned} x_1 + 10x_2 &= 11 \\ 10x_1 + 101x_2 &= 111 \end{aligned} \qquad (1)$$

$x_1 = x_2 = 1.$

Then considering the set

$$\begin{aligned} x_1 + 10x_2 &= 11 \\ 10.1x_1 + 100x_2 &= 111 \end{aligned} \qquad (2)$$

which is only a slight variation of the preceding set, but whose solution is $x_1 = 10$, $x_2 = 0.1$.

Systems (1) and (2) are simple systems which are ill-conditioned and can be observed as such. By inspection it is seen in both systems that the equations are almost dependent or, stated geometrically, they are nearly parallel. Thus, the intersection of two nearly parallel lines will change greatly if the coefficients are changed in the equations, this intersection being the solution of the two equations.

For larger systems parallelism refers to hyperplanes and in most cases is hard to recognize or hard to detect due to the increased number of equations.

Consider a system of equations written in matrix form

$$AX = B \tag{3}$$

where

$$A = \left[a_{ij}\right] = (n \times n) \text{ coefficient matrix}$$

$$X = \left[x_i\right] = (n \times 1) \text{ column matrix of unknowns}$$

$$B = \left[b_i\right] = (n \times 1) \text{ column matrix of constants}$$

If an approximate solution to (3) is x', then

$$AX' - B = R \qquad R = \left[r_i\right] \tag{4}$$

where R is an $(n \times 1)$ column matrix that measures how well the solution satisfies the original system (3). The quantities r_i are called residuals. Consider again the two equations

$$\begin{aligned} x_1 + 10x_2 &= 11 \\ 10x_1 + 101x_2 &= 111 \end{aligned} \tag{5}$$

If the solution is approximated by $x_1 = 1.0001$ and $x_2 = 1.0001$, the residuals r_i are $r_1 = .01$, $r_2 = .01$. But if the approximate solution $x_1 = 10.1$ and $x_2 = .09$ is substituted in the system, the residuals r_i are $r_1 = 0.0$, $r_2 = .01$. It would be natural to conclude that the solution $x_1 = 10.1$ and $x_2 = .09$ was nearer to the true solution than $x_1 = 1.0001$ and $x_2 = 1.0001$ which is, however, false since it is known that the true solution is $x_1 = x_2 = 1$.[2]

From the above it is shown that satisfying the equation by consideration of how small the residuals are does not guarantee that the solution is near the true solution, especially if the system is ill-conditioned.

Moulton[3] also illustrates the above behavior in the solution of the system:

$$\begin{aligned} .34622x + .35381y + .36518z &= .24561 \\ .89318x + .90274y + .91143z &= .62433 \\ .22431x + .23642y + .24375z &= .17145 \end{aligned} \tag{6}$$

found as

$$x = -1.02706 \qquad y = 2.09191 \qquad z = -.380476 \tag{7}$$

Upon substitution of this solution into (6) it is found to satisfy the equations to the last decimal place.
Also the solutions

$$\begin{aligned} x &= -1.022773 \qquad y = 2.084125 \qquad z = -0.376941 \\ x &= -1.031229 \qquad y = 2.099457 \qquad z = -0.383879 \end{aligned} \tag{8}$$

will satisfy (6) to the last decimal place. The variation in x is nearly one percent and in z nearly two percent. Therefore, assuming the equations (6) to be accurate to five decimal places, their solution is determinate only to two decimal places.

If equation (6) is written in the general form,

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned} \tag{9}$$

the expression for x by Cramer's Rule in the solution of simultaneous equations is

$$x = \frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}} = \frac{Dx}{D} \tag{10}$$

with corresponding expressions for y and z.

Suppose the numerical value of D is small. If a_i, b_i, and c_i are defined only to five decimal places and the first three decimal places in D are zeros, D is really given to only two decimal places. Consequently, x is defined only to two places. In the example above D is small, and this is the explanation of the variations in the solutions .

The above example shows how the value of the coefficient determinant, D, plays an important role in the solution of a system of linear equations. The smallness of the determinant of coefficients in the system is a measure proposed by several, but in this form the value is not too significant since the system can be multiplied by a constant without altering the solution and can make the determinant as large as desired.

A method of testing the condition of a system of equations is to find the determinant of the normalized coefficient matrix which is equal to the determinant of the coefficient matrix divided by

$$\prod_{i=1}^{n} \sqrt{\sum_{j=1}^{n} a_{ij}^{2}} \tag{11}$$

A comparison of this value is then made with $\pm$ 1.[2] The quantity obtained is the sine of the angle between two lines in a system of two equations and two unknowns. This is evident from the definition of the cross product of two vectors $\bar{A}$ and $\bar{B}$ that

$$|\bar{A} \otimes \bar{B}| = |\bar{A}|\ |\bar{B}| \sin\theta \tag{12}$$

or

$$\sin\theta = \frac{|\bar{A} \otimes \bar{B}|}{|\bar{A}|\ |\bar{B}|} \tag{13}$$

where the symbol $\otimes$ indicates cross product. For systems with three equations and three unknowns or more the physical interpretation is more complicated.

Booth[2] and Bodewig[4] have considered another measure of ill-conditioning in a system of equations. Considering the system of equations in matrix form AX = B, this measure is the ratio $| \lambda_{max} / \lambda_{min} |$, where λ_{max} is the largest characteristic value of the coefficient matrix A, and λ_{min} is the smallest. This measure or the ratio of $| \lambda_{max} / \lambda_{min} |$ can be shown to be the ratio of the greatest to least axis of a hyper-ellipsiod. The equation of such comes about by consideration of the following.

If, as before, the inaccuracy of an approximate solution x' is expressed in terms of residuals as

$$AX' - B = R \tag{14}$$

then for the i^{th} equation there is associated a residual r_i. Instead of n residuals to express the inaccuracy, there is needed a single number which will express the inaccuracy. The length of R is considered as this number. The length of R is defined as

$$R^2 = (R,R) = \sum_{i=1}^{n} r_i^2 \tag{15}$$

which in turn is

$$R^2 = ((AX' - B), (AX' - B)) \tag{16}$$

or writing the above equation as a quadric form in x_i

$$R^2 = \sum_{i=1}^{n} (a_{i1}x'_1 + a_{i2}x'_2 + \dots a_{in}x'_n - b_i)^2 \tag{17}$$

It can be shown that equation (17) can be transformed into a new set, Z_i, where the new set is a linear combination of only squares of the new set Z_i, the cross products being eliminated.

The transformation which will give the required results is an orthogonal transformation which will reduce the equation to the form

$$R^2 = \lambda_1 Z_1^2 + \lambda_2 Z_2^2 + \dots + \lambda_n Z_n^2 \tag{18}$$

where λ_i's are the characteristic values of the coefficient matrix A.

Equation (18) is now the equation of an hyper-ellipsoid where the semi-axis have length $\frac{1}{\sqrt{\lambda_i}}$. Thus, the ratio $\lambda_{max} / \lambda_{min}$ is the ratio of major to minor axis of this hyper-ellipsoid. Also, if any of the λ_i are small, large values of Z_i can be accompanied by small residuals. The ratio $|\lambda_{max}/\lambda_{min}|$ is a means of indicating this condition which is characteristic of ill-conditioned systems.[2] For the above geometrical interpretation, it is assumed that the matrix A is positive definite, meaning all $\lambda_i > 0$.

Other test measures for detection of ill-conditioned equations are given by Bodewig.[4] They are the N-number, M-number, and a quantity called μ which is the dominant term in the expansion of the determinant of the coefficient matrix by minors divided by the determinant of the coefficient matrix.

If the diagonal element happens to be the dominant term in the determinant of the coefficient matrix A then μ is given as

$$\mu = | a_{11} \cdot a_{22} \cdot a_{33} \cdots a_{nn}| / |A| \tag{19}$$

In ill-conditioned systems μ could be large in magnitude as will be shown in the investigations to be considered later. If $|A|$ is small, then it is expected that μ would would be large.

The N-number is a measure of ill-conditioning based upon the coefficient matrix A, the inverse of A, and the order n of the system. Stated symbolically, the N-number is given as

$$\text{N - number} = \frac{1}{n} N(A) \cdot N(A^{-1}) \tag{20}$$

where terms $N(A)$ and $N(A^{-1})$ are called the norms of A and A^{-1} which is defined as

$$N(A) \pm \sqrt{\sum_{j=1}^{n} \sum_{i=1}^{n} a^{2}_{ij}} \tag{21}$$

where a_{ij} are the elements of the n x n coefficient matrix A. $N(A^{-1})$ is similar except the summation will be on the elements of A^{-1}.

Since by definition

$$A^{-1} = \frac{\text{adj } A}{|A|} \tag{22}$$

and if A is near-singular, then $|A|$ will be small and A^{-1} will have some elements which are large in magnitude. Hence $N(A^{-1})$ will very likely be large, and this large number multiplied by $N(A)$ divided by (n) will result in a large N-number. Thus, this measure might indicate when a system of equations is ill-conditioned.

The M-number stated symbolically is

$$\text{M-number} = n \cdot m\,(A) \cdot m\,(A^{-1}) \tag{23}$$

where $m(A)$ and $m(A^{-1})$ designate the maximum element of matrix A and A^{-1} respectively. Thus, as in the preceding measure if A is near-singular, then $|A|$ will be small; and the largest element of A^{-1} in magnitude could be large depending, of course, upon the original elements of A and the smallness of $|A|$. If the system is ill-conditioned, then it seems reasonable that the product $n \cdot m\,(A) \cdot m\,(A^{-1})$ would be large.

The subject of noise in the solution of large linear system has received attention by C. Lanczos.[5]

A method for solving ill-conditioned equations is given by E. Bodewig,[4] N. Macon,[1] and K. Eisemann.[6]

Similar conditions for detection of ill-conditioning and an example is given by J. Todd.[7]

Nowhere in the available literature was the attempt made to find the most suitable measure of ill-conditioning.

III. DISCUSSION

When the determinant of the matrix of coefficients in a set of n linear non-homogeneous equations in n unknowns is not zero, then that set has a unique solution. This, however, does not take into account the practical point of view. If the coefficients are furnished by observations, of measurements, they are not exact but will only be an approximation to a certain number of decimal places. Practical considerations or devices used for measurements have, unfortunately, physical limitations.

The mere mathematical solution of a set of ill-conditioned equations frequently hides or overshadows the dangers which arise when physical noise, caused by the inexactitude of the measurements, are introduced in the system.

The author attempted to devise a measure of ill-conditioning which he hopes might be worthy of mention.

Consider again the system of equations

$$AX = B \tag{24}$$

where A is the coefficient matrix, X the unknown column matrix, and B a column matrix of constants.

If the coefficients and the right-hand members of (24) are only approximations, then the true values of the coefficients could be represented as $A + \Delta A$, and the true values of the right-hand members as $B + \Delta B$, where ΔA and ΔB are the errors of approximation associated with the coefficient matrix A, and the matrix B respectively. Thus, the true system would be of the form

$$(A + \Delta A) \quad (X + \Delta X) = (B + \Delta B) \tag{25}$$

where $X + \Delta X$ is the true solution. Subtracting equation (24) from equation (25), there results[8]

$$A \cdot \Delta X + \Delta A \cdot X + \Delta A \cdot \Delta X = \Delta B \tag{26}$$

or

$$A \cdot \Delta X = \Delta B - (\Delta A \cdot X + \Delta A \cdot \Delta X) \tag{27}$$

If the system is well-behaved then for a small change ΔA, there would be expected small changes in ΔX so that if all elements of the product $\Delta X \cdot \Delta A$ are small and can be neglected, we can justify writing equation (27) as

$$A \cdot \Delta X \approx \Delta B - (\Delta A \cdot X) \tag{28}$$

If all the Δa_{ij} and Δb_{ij}, the possible errors or bounds on the errors of approximation are known, the approximate errors or deviations of Δx_i caused by this change or possible errors can be found.

In practice ΔA and ΔB are usually known such that the errors in each do not exceed a certain magnitude, ϵ, or

$$-\epsilon \leq \Delta a_{ij} \leq \epsilon \qquad -\epsilon \leq \Delta b_i \leq \epsilon \tag{29}$$

Let the right-hand members of system (28) be replaced by the matrix N where

$$N = \Delta B - (\Delta A \cdot X) \qquad N = \left[\eta_i \right] \tag{30}$$

giving

$$A \cdot \Delta X = N \tag{31}$$

where X in equation (30) is the solution of the system $AX = B$. If the errors ΔA and ΔB are such that their magnitude does not exceed ϵ, then

$$| \eta_i | \leq E \tag{32}$$

where $E = (1 + |X_1| + |X_2| + . . . + |X_n|)\epsilon$ or the maximum value possible in equation (30).

Comparison of the two systems $A \cdot \Delta X = N$ and $AX = B$ indicates that they are similar in the fact that both contain the coefficient matrix A. The augmented matrices of both systems are such that only the last columns are different as shown below

$$\begin{array}{cc} AX = B & A\Delta X = N \\ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} & b_n \end{bmatrix} & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & \eta_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & \eta_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} & \eta_n \end{bmatrix} \end{array} \tag{33}$$

The Gauss-Jordan elimination method for solving a system of linear equations can be applied equally well to solve systems with more than one B column, or any number of constant columns at the same time. Since it is assumed that each element η_i of matrix N is such that $| \eta_i | \leq E$ system (31) is represented by the augmented matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & \pm & E \\ a_{21} & a_{22} & \cdots & a_{2n} & \pm & E \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdot & \cdots & a_{nn} & \pm & E \end{bmatrix} \tag{34}$$

If the last column of (34) is included as an extra constant column in the solution of the AX = B in a Gauss-Jordan elimination, the augmented matrix would be of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \pm & E \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \pm & E \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \pm & E \end{bmatrix} \tag{35}$$

where both systems $A \cdot \Delta X = N$ and $AX = B$ are represented.

The Gauss-Jordan solution of (35) is carried out such that the results of the operations on the last column are always positive. This is accomplished by replacing subtraction by additions and working with the absolute value of the elements in the last column. The result of solving by this method would give the solution X in the n + 1 column and the largest ΔX that could occur in the last column. This would be some multiple of E. This is not an upper bound since the product $\Delta A \cdot \Delta X$ was neglected.

The author instead of using some value of E for the elements of the last column, used the column with unity as each element and studied what might be thought of as multiples of some error in the last column.

As was shown earlier in the system

$$\begin{aligned} x_1 + 10x_2 &= 11 \\ 10x_1 + 101x_2 &= 111 \end{aligned} \tag{37}$$

If the coefficients are changed by approximately one percent in the second equation, large difference in solutions result. This equation is known to be ill-conditioned because of its sensitivity to the coefficients and its evident parallelism in equations.

The Gauss-Jordan method applied to (37) is the augmented matrix.

$$\begin{bmatrix} 1 & 10 & 11 \\ 10 & 101 & 111 \end{bmatrix} \tag{38}$$

Now if equation (38) is modified somewhat by adding an extra column with the value of unity as its elements or

$$\begin{bmatrix} 1 & 10 & 11 & 1 \\ 10 & 101 & 111 & 1 \end{bmatrix} \tag{39}$$

The solution represented after a number of operations on the rows and always adding or dividing by absolute values in the last column is

$$\begin{bmatrix} 1 & 0 & 1 & 111 \\ 0 & 1 & 1 & 11 \end{bmatrix} \tag{40}$$

where the last column indicates the multiples of some error $|E|$ in each equation. It is noticed that this is a fairly large multiple.

In connection with the preceding discussion, a study was made of the system

$$\begin{aligned} x_1 + 10x_2 &= 11 \\ x_1 + mx_2 &= 11.1 \end{aligned} \qquad (41)$$

where

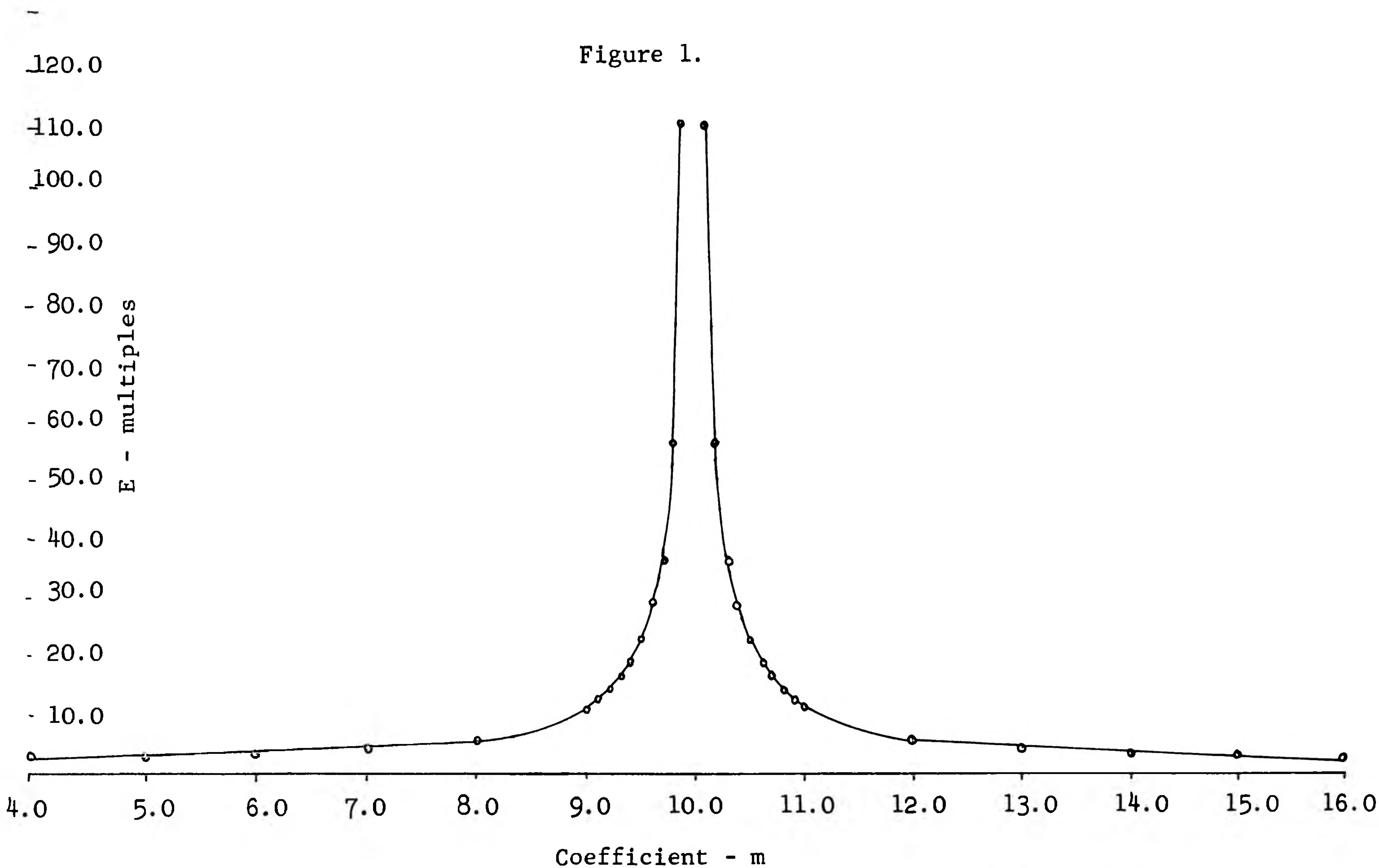
$$m = 3, 4, 5, \ldots 9, 9.1, 9.2, \ldots 9.9, \\ 10.1, 10.2, \ldots 10.9, 11, 12, \ldots 15$$

to observe the effect of changing the coefficient m in relationship with the average E-multiple obtained by the method described above. The results of this study are shown graphically in Fig. I. The graph clearly indicates the parallelism of the equations when m approaches the value of 10. The value of 10 was excluded since the system would not possess a unique solution in that case. The study shows that there is a relationship between the coefficient and E-multiples which is some indication of ill-conditioning. There is needed additional study of larger and different systems, however, before anything definite can be said about whether or not the E-multiples indicate ill-conditioning.

A program was written for the IBM 1620 digital computer to calculate the values of the E-multiples for the different values of m. In effect, the program was a Gauss-Jordan elimination with only slight modification.

An investigation was made of several systems of equations as to ill-conditioning, applying various tests and measures as given in the literature along with the author's measure.

Figure 1.



120.0
110.0
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
E - multiples
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
12.0
13.0
14.0
15.0
16.0
Coefficient - m

In applying the test measures, two programs were written in the Fortran Computer language for the IBM 1620 digital computer at the School of Mines Computer Center, and two standard programs were used from files in the Computer Center.

Program I consisted of a Gauss-Jordan with pivot picker, the determinant of the coefficient matrix, the determinant of the normalized coefficient matrix, the product of the diagonal elements divided by the determinant of the coefficient matrix, and the author's measure.

In Program II the inverse of the coefficient matrix was determined, and the measure's N-number and M-number were found. This program called for the subroutine MINVRT(A,N) on file in the School of Mines and Metallurgy Computer Center to calculate the inverse of the coefficient Matrix A of order N.

The program JACOBI was also used from the Computer Center files to find the eigenvalues of the coefficient matrix. This program uses the Jacobi iterative method to find all the eigenvalues of a real symmetric matrix.

When one is confronted with finding the eigenvalues or characteristic values of a matrix, consideration is given to the solution of a homogenous set of equation of the form

$$AX - \lambda X = 0 \tag{42}$$

It is seen that this possesses a solution other than the trivial solution if and only if the determinant of the coefficient matrix vanishes or

$$|A - \lambda I| = 0 \tag{43}$$

This says that λ must be a root of a polynomial equation of degree n, where n is the order of the matrix A. The polynomial is called a characteristic polynomial, and the λ's are called the eigenvalues or characteristic values of the matrix A. When the coefficient matrix of the systems investigated were not symmetric, this approach was used to find the eigenvalues. The characteristic polynomial equation was formed, and a standard program on file in the School of Mines and Metallurgy Computer Center was used tc obtain the roots which are the eigenvalues of the coefficient matrix.

On the systems tested, the N-number ranged from about 1.6 to 6.6 for systems which were well-behaved and 38 to 8,090 for ill-conditioned systems. The N-number is a good indication of ill-conditioning, however, in practice it requires too many operations for its determination. The one big disadvantage is the calculation of the inverse of the coefficient matrix. In large ill-conditioned systems the calculation of the inverse is a serious problem in that cumulative effect of round-off is likely to occur.

The M-number has the same disadvantage in that of determining the inverse of the coefficient matrix. The M-number gave a good measure of ill-conditioning in the systems investigated, but it would not be practical unless the system was solved by inversion in the first place. Of course, as stated above, this is a serious problem in itself if the system is ill-conditioned. The values of the M-number ranged from 10.0 to 50.0 for systems that were well-behaved and 50.0 to 58,000 for ill-conditioned systems. Both the M-number and N-number were large numbers when the system was ill-conditioned.

The measure μ, or the dominant term in the expansion of the determinant of the coefficient matrix (see page 9) divided by the value of the determinant, was found to give a good measure of ill-conditioning. The values obtained from the systems investigated ranged between 0.0 to 3.0 in magnitude for the well-behaved systems and from 3.0 to 57,600 in magnitude for ill-conditioned systems. In small systems, such as the ones investigated, the calculation of μ was fairly easy since the dominant term could be determined by inspection. In larger systems, however, the dominant term would be hard to single out, thus creating a problem in determining μ. The Gauss-Jordan method with pivot picker provides an easy method for calculation of the determinant of the coefficient matrix. Even so, the dominant term of this determinant still would need determining. If another method of solution is used, a separate calculation for the evaluation of the dominant term and the full value of the determinant of coefficients would have to be made. This would not be practical in most cases.

Since it required no difficulty in programming, the product of the diagonal elements were computed as the dominant term of the determinant of coefficients. Several systems investigated were of this character. However, when the dominant term was not the product of the diagonal elements, the calculation of μ was made by observing the dominant term and dividing by the value of the determinant of the coefficients. The determinant was evaluated by program I.

The ratio $|\lambda_{max}/\lambda_{min}|$ was found to be a very good indication of ill-conditioning in most of the cases tested. The exception being when the coefficient matrix was not symmetric. System 4 was of this nature where two of the characteristic values of the coefficient matrix was complex. Thus, this measure is not significant when some of the characteristic values are complex. In other systems considered, where the coefficient matrix was not symmetric, the characteristic values were found to ratio $|\lambda_{max}/\lambda_{min}|$ was still an indication of ill-conditioning. The values ranged from 4.0 to 20.0 for systems believed to be well-behaved and 54.0 to 85,000 for ill-conditioned systems.

Even though apparently good results are obtained using this measure, the complexity of calculating $|\lambda_{max}/\lambda_{min}|$ is such that this measure is of little practical use. The calculation of λ_{max} and λ_{min} is far more complex than that of solving the original system. A measure is needed which will take the least amount of time and calculation to be of practical use.

The determinant of the normalized coefficient matrix, designated by |A-normalized|, seems to be the better of the methods discussed so far. This measure is more practical than the others in that it requires less operations to calculate. The values obtained for well-behaved systems were in the range .2 to .6, whereas for ill-conditioned systems the values ranged from 8×10^{-8} to .04. This measure has the advantage in that it requires nothing more than the coefficient matrix and simple operations on it, whereas the other measures require both the original and the inverse coefficient matrix. This measure gives good indication of ill-condition and is simple to calculate.

The author's measure seemed to indicate ill-conditioning for systems whose elements or coefficients were small in magnitude. However, when the test was made on systems whose elements were large, the measure failed to indicate ill-conditioning. Three systems were noted of this character, and they were all three constructed such that they were ill-conditioned. All three had elements which were large in magnitude.

These results indicate that further study was needed. It was found that the last column in the augmented solution of $AX = B$ and $A\Delta X = N$, assuming that the column was originally all elements of unity, is a maximum multiple of E that could occur for that system. If these multiples are denoted by m_i in the last column, the product m_i E would represent a bound on the deviation of the solution or

$$\Delta x_i = m_i E \tag{44}$$

where the value of E can be calculated from the relation $E = (1 + |x_1| + |x_2| + . . . |x_n|)\ \epsilon$, for a change of e magnitude in all coefficients and constant terms of the system. Thus equation (44) is written in the form

$$\Delta x_i = m_i \quad (1 + |x_1| + |x_2| + . . . x_n)\ \epsilon \qquad (45)$$

Comparison of Δx_i and x_i should be some indication of ill-conditioning realizing that large change in solution accompanied by small change in coefficients is characteristic of ill-conditioning.

It may be known that the coefficients are accurate to a number of places. If this is the case, a value of ϵ may be used corresponding to the measure of doubt in the coefficients. If it is not known how accurate the coefficients are, a value of ϵ can be chosen to test the sensitivity of the solution. The value assumed for ϵ should be a reasonable value associated with the particular system.

A study was made of changes in solution, Δx for each system investigated. A value of ϵ was chosen for each particular system, and the results or the change in solution Δx tabulated in the last column of Table I.

The results of this study give more of an indication of ill-conditioning than just the E-multiples considered previously. Where the E-multiples failed to indicate ill-conditioning when the coefficients of the systems studied were large, this measure indicated ill-conditioning satisfactorily.

To illustrate the sensitivity of the coefficients in the case of an ill-conditioned system of equations, the coefficients of the systems investigated were changed slightly and the corresponding solution noted.

The result of changing or deviating only the coefficients in the last equation of each system caused the solutions to change greatly. The original and deviated systems are shown with their corresponding solutions in Appendix I.

A solution of the various systems was carried out with the Gauss-Jordan elimination method without the feature of picking the largest pivot to note any change in solutions as compared with pivot picker. The solutions obtained were the same. If larger systems were considered, this would make a difference in the solution; but since it was only small systems considered, the change was not significant.

TABLE I

MEASURES AND VALUES FOR SYSTEMS STUDIED

System	N-number	M-number	$\lvert A \rvert$	μ	$\lvert A_{norm} \rvert$	$\lvert \lambda_{max} / \lambda_{min} \rvert$	Max. E-mult.	ϵ	ΔX
1	752.4	2,720.	1.0	5,000.	.00002	3,025.0	190.4	.1	95.20 57.61 22.65 13.07
2	5,200.	20,400.	1.0	101.	.00098	10,407.	111.2	.1	33.36 3.30
3	149.6	270.	1.0	512.	.00039	448.	37.0	.1	2.40 14.80 14.00
4	240.0	704.	.00005	1,478.4	.000132	complex	615.09	.0005	.82098 1.69149 .87384
5	38.33	63.0	.19	11.26	.042	57.	21.42	.05	3.437 2.237
6	3,900.00	25,900.0	1.7×10^{-7}	57,602.	.0000010	15,500	29,905.	.00005	.2895 3.1650 6.4740 4.7950

TABLE I (CONT.)

System	N-number	M-number	$\lvert A \rvert$	μ	$\lvert A_{norm} \rvert$	$\lvert \lambda_{max}/\lambda_{min} \rvert$	Max. E-mult.	ϵ	ΔX
7	1,709.7	16,100	1.0	16,800	8×10^{-8}	85,180	415.0	.01	8.34 23.04 24.90 12.60 2.52
8	8,087.4	58,067.	.3735	60.7	.0044	24,260	161.4	.01	.1224 6.4540 .0610
9	72.38	270.	.0045	48.	.00767	208.	169.3	.01	4.592 .106 6.773
10	2.58	10.73	595.	2.42	.2439	7.58	.656	0.10	.328 .261 .322 .278
11	6.58	52.18	1,104.	1.2	.587	18.4	.42	0.10	.0231 .0811 .1681
12	2.44	16.86	1,602,556.0	1.86	.2883	7.10	.0919	0.10	.0264 .0459 .0441 .0270

TABLE I (CONT.)

System	N-number	M-number	$\lvert A \rvert$	μ	$\lvert A_{norm} \rvert$	$\lvert \lambda_{max}/\lambda_{min} \rvert$	Max. E-mult.	ϵ	ΔX
13	1.662	10.7	1,389.	.967	.5587	4.0	.501	.01	.0313 .0294 .0334 .0501
14	2.578	12.145	548.2	.4652	.2129	6.0	3.02	.01	.1026 .1536 .2114 .1463
15	1,666,447.	6,100,830.	3,457,498.	9.7600	$.5x10^{-12}$	83.0	5.0468	0.10	95.62400 151.90800 .00078
16	7,667.	42,640.	207,987.	50,604.0	.000002	23,000.0	2.400	0.5	2.416 4.800 .067
17	299.9	1,280.	171,947.	113.99	.00321	764.0	1.500	0.5	1.500 .013 3.000
18	17.72	75.	.96	25.00	.00696	54.0	10.416	0.1	1.430 4.370 5.210

IV. CONCLUSIONS

From the results obtained in the investigation of measures of ill-conditioning, it is believed that the most suitable measure considered was that of the determinant of the normalized coefficient matrix. This measure was found to indicate ill-conditioning quite well. It was found that, with any of the measures considered, a considerable amount of computation is required of each. The determinant of the normalized coefficient matrix takes considerable amount of computation unless the solution of the system is carried out by Gauss-Jordan elimination with pivot picker. The determinant in this case is the product of the pivot elements times $(-1)^p$ where p is the number of times the rows or equations of the system are interchanged. Using the number of multiplications as a measure of efficiency and n as the order of the coefficient matrix, there are approximately $n^2 + 2n$ multiplications to perform in calculating the determinant of the normalized coefficient matrix when the Gauss-Jordan algorithm with pivot picker is used. Generally the evaluation of a determinant requires on the order of $\frac{n^3}{3}$ multiplications. Even so, it still would be desirable that programs for the solution of linear equations include the computation of the determinant.

Some of the other measures studied, such as the N-number, required the calculation of the inverse of the coefficient matrix. The calculation of an inverse matrix requires on the order of n^3 multiplications. This amount plus additional operations on the elements of the inverse and original coefficient matrix far exceeds the amount of operations when the determinant of the normalized coefficient matrix in conjunction with the Gauss-Jordan algorithm is applied.

The measure μ is an attractive measure in that it requires about $2n + 1$ multiplications when the Gauss-Jordan algorithm is used. It has the disadvantage due to the determination of the dominant term in the expansion of the determinant of the coefficient matrix by minors. If it is assumed, in all systems, that the dominant term is the product of the diagonal elements, the measure is found to give poor indication of ill-conditioning. The above assumption would simplify calculation of the measure greatly; however, system 15 of this study gave a number which indicated a well-behaved system when the product of the diagonal elements were used as the dominant term in the expansion of the determinant of the coefficient matrix. The system was ill-conditioned by construction. Thus, no prescribed choice for the dominant term will work for all systems.

The author's measure requires about n^2 multiplications to calculate. It must also be used in conjunction with the Gauss-Jordan algorithm to be very useful. It could be a useful measure in other respects as well as a measure to indicate ill-conditioning. For reasonable choices of ϵ

the measure seems to indicate ill-conditioning satisfactorily. If the systems are well-behaved, the values obtained are still worth while in that they give a bound on the errors in solutions.

V. SUMMARY

Ill-conditioning is a problem that can be easily overlooked when one is confronted with obtaining the solution of a system of non-homogeneous linear equations. The result of overlooking this problem may lead to solutions which are unreasonable. A study was made of several measures which indicated ill-conditioning, and these measures were applied to a number of different systems of equations. The largest system of equations considered was a system of five equations in five unknowns. Programs were written for the IBM 1620 Digital Computer to calculate the various measures considered. The values obtained were then tabulated and studied in an effort to determine which measure (1) best indicated ill-conditioning, (2) was the simpliest to compute, and (3) required the least time to perform the computations.

Systems of equations of order greater than five were not studied since the same measures can be applied to systems of any order. It is also reasonable to assume that the most suitable measure for small systems would also be the most desirable measure for large systems.

APPENDIX I

AUGMENTED SYSTEMS AND THEIR SOLUTIONS

System 1

Original

$$\begin{bmatrix} 5 & 7 & 6 & 5 & 23 \\ 7 & 10 & 8 & 7 & 32 \\ 6 & 8 & 10 & 9 & 33 \\ 5 & 7 & 9 & 10 & 31 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 1.0 \\ 1.0 \\ 1.0 \end{bmatrix}$$

Deviated

$$\begin{bmatrix} 5 & 7 & 6 & 5 & 23 \\ 7 & 10 & 8 & 7 & 32 \\ 6 & 8 & 10 & 9 & 33 \\ 5.1 & 7.1 & 9.1 & 10.1 & 31 \end{bmatrix}$$

$$X = \begin{bmatrix} -2.0769 \\ 2.8461 \\ 1.9230 \\ .3846 \end{bmatrix}$$

System 2

$$\begin{bmatrix} 1 & 10 & 11 \\ 10 & 101 & 111 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 1.0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 10 & 11 \\ 10.1 & 100 & 111 \end{bmatrix}$$

$$X = \begin{bmatrix} 10.0 \\ 0.1 \end{bmatrix}$$

System 3

$$\begin{bmatrix} 9 & 9 & 8 & 26 \\ 9 & 8 & 7 & 24 \\ 8 & 7 & 6 & 21 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 1.0 \\ 1.0 \end{bmatrix}$$

$$\begin{bmatrix} 9 & 9 & 8 & 26 \\ 9 & 8 & 7 & 24 \\ 8.1 & 7.1 & 6.1 & 21 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.333 \\ -2.000 \\ 4.000 \end{bmatrix}$$

System 4

$$\begin{bmatrix} .34622 & .35381 & .36518 & .24561 \\ .89318 & .90274 & .91143 & .62433 \\ .22431 & .23642 & .24375 & .17145 \end{bmatrix}$$

$$X = \begin{bmatrix} -1.0270 \\ 2.0919 \\ -0.3804 \end{bmatrix}$$

$$\begin{bmatrix} .34622 & .35381 & .36518 & .24561 \\ .89318 & .90274 & .91143 & .62433 \\ .22655 & .23406 & .24619 & .17145 \end{bmatrix}$$

$$X = \begin{bmatrix} -31.8250 \\ 47.5685 \\ -15.2422 \end{bmatrix}$$

System 5

$$\begin{bmatrix} 2.00 & 3.00 & 5.00 \\ .65 & 1.07 & 1.68 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.6317 \\ .5789 \end{bmatrix}$$

$$\begin{bmatrix} 2.00 & 3.00 & 5.00 \\ .66 & 1.05 & 1.68 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.75 \\ .50 \end{bmatrix}$$

System 6

$$\begin{bmatrix} 1.00000000 & .500000000 & .333333333 & .250000000 & 2.083333333 \\ .50000000 & .333333333 & .250000000 & .200000000 & 1.283333333 \\ .33333333 & .250000000 & .200000000 & .166666666 & .949999999 \\ .25000000 & .200000000 & .166666666 & .142857142 & .759523808 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.000001 \\ .999895 \\ 1.000240 \\ .999848 \end{bmatrix}$$

Change of last row of:

.25300000	.203000000	.169666666	.142587142	.759523808

$$X = \begin{bmatrix} 1.101620 \\ -\ .219380 \\ 4.048450 \\ -1.032290 \end{bmatrix}$$

System 7

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 5 \\ 1 & 2 & 3 & 4 & 5 & 15 \\ 1 & 3 & 6 & 10 & 15 & 35 \\ 1 & 4 & 10 & 20 & 35 & 70 \\ 1 & 5 & 15 & 35 & 70 & 126 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 5 \\ 1 & 2 & 3 & 4 & 5 & 15 \\ 1 & 3 & 6 & 10 & 15 & 35 \\ 1 & 4 & 10 & 20 & 35 & 70 \\ 1.0 & 5.1 & 15.1 & 35.2 & 70.3 & 126 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 1.0 \\ 1.0 \\ 1.0 \\ 1.0 \end{bmatrix} \qquad X = \begin{bmatrix} 2.4274 \times 10^{-5} \\ 4.99994 \\ -4.99995 \\ 4.99999 \\ 1.07497 \times 10^{-6} \end{bmatrix}$$

System 8

$$\begin{bmatrix} 30.0 & .00325 & 60.5 & 90.50325 \\ .00325 & .00620 & .00620 & .01565 \\ 60.5 & .00625 & 120.0 & 180.50625 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 1.0 \\ 1.0 \end{bmatrix} \qquad \begin{bmatrix} 30.0 & .00325 & 60.5 & 90.50325 \\ .00325 & .00620 & .00620 & .01565 \\ 61.105 & .00618 & 121.2 & 180.50625 \end{bmatrix}$$

$$X = \begin{bmatrix} -.79445 \\ 1.05083 \\ 1.88981 \end{bmatrix}$$

System 9

$$\begin{bmatrix} .90 & .30 & .60 & 1.80 \\ .30 & .60 & .20 & 1.10 \\ .60 & .21 & .39 & 1.20 \end{bmatrix} \qquad \begin{bmatrix} .90 & .30 & .60 & 1.80 \\ .30 & .60 & .20 & 1.10 \\ .61 & .22 & .40 & 1.20 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 1.0 \\ 1.0 \end{bmatrix} \qquad X = \begin{bmatrix} -2.00 \\ 1.00 \\ 5.50 \end{bmatrix}$$

System 10

$$\begin{bmatrix} 8 & -1 & 5 & 1 & 13 \\ -1 & 5 & 2 & -1 & 5 \\ 5 & 2 & 9 & -1 & 15 \\ 1 & -1 & -1 & 4 & 3 \end{bmatrix} \qquad \begin{bmatrix} 8 & -1 & 5 & 1 & 13 \\ -1 & 5 & 2 & -1 & 5 \\ 5.05 & 1.98 & 9.09 & -1.01 & 15 \\ 1.01 & -1.01 & -.99 & 3.96 & 3 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 1.0 \\ 1.0 \\ 1.0 \end{bmatrix} \qquad X = \begin{bmatrix} 1.01664 \\ 1.01311 \\ .97587 \\ 1.00064 \end{bmatrix}$$

System 11

$$\begin{bmatrix} 49.0 & -7.0 & 1.0 & 43.0 \\ -7.0 & 9.0 & 1.0 & 3.0 \\ 1.0 & 1.0 & 3.0 & 5.0 \end{bmatrix} \qquad \begin{bmatrix} 49.0 & -7.0 & 1.0 & 43.0 \\ -7.0 & 9.0 & 1.0 & 3.0 \\ 1.01 & 1.01 & 3.03 & 5.0 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 1.0 \\ 1.0 \end{bmatrix} \qquad X = \begin{bmatrix} 1.00072 \\ 1.00251 \\ .98242 \end{bmatrix}$$

System 12

$$\begin{bmatrix} 69.0 & 4.0 & 12.0 & 30.0 & 115.0 \\ 4.0 & 36.0 & 8.0 & 16.0 & 64.0 \\ 12.0 & 8.0 & 20.0 & 7.0 & 47.0 \\ 30.0 & 20.0 & 7.0 & 60.0 & 117.0 \end{bmatrix} \qquad \begin{bmatrix} 69.0 & 4.0 & 12.0 & 30.0 & 115.0 \\ 4.0 & 36.0 & 8.0 & 16.0 & 64.0 \\ 12.0 & 8.0 & 20.0 & 7.0 & 47.0 \\ 30.1 & 20.1 & 7.1 & 60.2 & 117.0 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 1.0 \\ 1.0 \\ 1.0 \end{bmatrix} \qquad X = \begin{bmatrix} 1.0053 \\ 1.0052 \\ .99913 \\ .98739 \end{bmatrix}$$

System 13

$$\begin{bmatrix} 6 & 3 & 0 & 0 & 12 \\ 3 & -7 & 3 & 0 & -2 \\ 0 & 3 & 8 & 3 & 42 \\ 0 & 0 & 3 & 4 & 25 \end{bmatrix} \qquad \begin{bmatrix} 6 & 3 & 0 & 0 & 12 \\ 3 & -7 & 3 & 0 & -2 \\ 0 & 3 & 8 & 3 & 42 \\ 0 & 0 & 3.03 & 3.96 & 25 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 2.0 \\ 3.0 \\ 4.0 \end{bmatrix} \qquad X = \begin{bmatrix} 1.00138 \\ 1.99723 \\ 2.99216 \\ 4.02368 \end{bmatrix}$$

System 14

$$\begin{bmatrix} 10.2 & -4 & 0 & 0 & 6.2 \\ -4 & 5 & -4 & 0 & -7.0 \\ 0 & -4 & 5 & -4 & -2.0 \\ 0 & 0 & -4 & -1 & -10.0 \end{bmatrix} \qquad \begin{bmatrix} 10.2 & -4 & 0 & 0 & 6.2 \\ -4 & 5 & -4 & 0 & -7.0 \\ 0 & -4 & 5 & -4 & -2.0 \\ 0 & 0 & -4.1 & -1.1 & -10.0 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.0 \\ 1.0 \\ 2.0 \\ 2.0 \end{bmatrix} \qquad X = \begin{bmatrix} .956387 \\ .888787 \\ 1.904600 \\ 1.991960 \end{bmatrix}$$

System 15

$$\begin{bmatrix} 11.5 & 4.5 & -2.0\text{x}10^6 & 50.0 \\ 10.0 & 1.0 & -1.25\text{x}10^6 & -50.0 \\ 6.0 & .5 & -.63\text{x}10^6 & 70.0 \end{bmatrix} \qquad \begin{bmatrix} 11.5 & 4.5 & -2.00\text{x}10^6 & 50.0 \\ 10.0 & 1.0 & -1.25\text{x}10^6 & -50.0 \\ 6.1 & .55 & -.635\text{x}10^6 & 70. \end{bmatrix}$$

$$X = \begin{bmatrix} 100.0 \\ 200.0 \\ .001 \end{bmatrix} \qquad X = \begin{bmatrix} -0.2205 \\ 44.4854 \\ .00007382 \end{bmatrix}$$

System 16

$$\begin{bmatrix} 8764.5 & 4382.0 & 2191.5 & 15338.0 \\ 4382.0 & 2191.5 & 1095.7 & 7669.2 \\ 2191.5 & 1095.7 & 510.0 & 3797.2 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 8764.5 & 4382.0 & 2191.5 & 15338.0 \\ 4382.0 & 2191.5 & 1095.7 & 7669.2 \\ 2191.0 & 1095.2 & 510.5 & 3797.2 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.00337 \\ 0.99994 \\ .98661 \end{bmatrix}$$

System 17

$$\begin{bmatrix} 428.0 & 108.0 & 214.0 & 750.0 \\ 108.0 & 428.0 & 55.0 & 591.0 \\ 214.0 & 55.0 & 106.0 & 375.0 \end{bmatrix} \qquad \begin{bmatrix} 428.0 & 108.0 & 214.0 & 750.0 \\ 108.0 & 428.0 & 55.0 & 591.0 \\ 214.5 & 55.5 & 106.5 & 375.5 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad X = \begin{bmatrix} .337195 \\ .996688 \\ 2.327280 \end{bmatrix}$$

System 18

$$\begin{bmatrix} 6 & 4 & 3 & 19 \\ 4 & 2 & 1.6 & 11.6 \\ 3 & 1.6 & 1.5 & 9.1 \end{bmatrix} \qquad \begin{bmatrix} 6 & 4 & 3 & 19 \\ 4 & 2 & 1.6 & 11.6 \\ 3.1 & 1.7 & 1.6 & 9.1 \end{bmatrix}$$

$$X = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \qquad X = \begin{bmatrix} 2.14815 \\ 1.88889 \\ -0.481481 \end{bmatrix}$$

BIBLIOGRAPHY

1. Macon, N., Numerical Analysis, 1st Ed. N. Y., John Wiley and Sons, 1963, pp. 65-69.

2. Booth, A., Numerical Methods, 2nd Ed., N. Y., Academic Press, Inc., 1957, pp. 80-85.

3. Moulton, F., American Mathematical Monthly, Vol. 20, 1913, pp. 242-249.

4. Bodewig, E., Matrix Calculus, 2nd Ed., N. Y., North Holland Publishing Company, 1959, pp. 133-137.

5. Lanczos, C., Applied Analysis, 1st Ed., N. J., Prentice Hall, Inc., 1956, pp. 167-170.

6. Eisemann, K., Quarterly of Applied Mathematics, Vol. 15-16, 1957-59, pp. 225-230.

7. Todd, J., Survey of Numerical Analysis, 1st Ed., N. Y., McGraw-Hill, 1962, pp. 239-243.

8. Hildebrand, Introduction to Numerical Analysis, 1st Ed., N. Y. McGraw-Hill, 1956, pp. 436-439.

9. Arden, B., An Introduction to Digital Computing, 1st Ed., Massachusetts, Addison-Wesley, Inc., 1963, pp. 110-111, 246-248.

VITA

The author was born September 9, 1938, at Verona, Missouri, the fifth son of Mr. and Mrs. Marion Elbe Calton.

His early education was received at Lone Hill Missouri and high school at Verona, Missouri. He entered Southwest Missouri State College in March 1958 and graduated in May 1962 with a B. S. in Mathematics.

In September 1962 the author entered the University of Missouri, School of Mines and Metallurgy for graduate work. From September 1962 to June 1963 he served as a graduate assistant in the Computer Center.