1
UNITED STATES ATOMIC ENERGY COMMISSION

K-731

**AN ITERATIVE METHOD FOR SOLVING SYSTEMS
OF LINEAR EQUATIONS**

**By
A. de la Garza**

February 15, 1951

**Carbide and Carbon Chemicals Company
UCC, K-25**

Technical Information Service, Oak Ridge, Tennessee

ABSTRACT

An iterative method for solving systems of linear equations is presented. The basis for the method is a systematic reduction of the sum of squares of residuals by which trial solutions may be improved to a desired accuracy. Remarks on application of the method are included.

PHYSICS

Reproduced direct from copy
as submitted to this office.

Work performed under
Contract No. W-7405-eng-26

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

AN ITERATIVE METHOD FOR SOLVING SYSTEMS OF LINEAR EQUATIONS

By A. de la Garza

Introduction

Consider the linear system

$$A\xi + \rho = \eta, \quad (1)$$

where we have given the matrix $A = (a_{ij})$, $i = 1(1)n$, $j = 1(1)m$, $n \geq m$, with rank m , and the column vector $\eta = (y_i)$. It is clear that we may let the column vector $\xi = (x_j)$ take any value we please and determine the residual column vector $\rho = (r_i)$ by difference.

Let us define the column vector $\hat{\xi} = (\hat{x}_j)$ to be the solution of (1) if for $\xi = \hat{\xi}$, the sum of squares of residuals.

$$\rho'\rho = (\eta - A\xi)'(\eta - A\xi)$$

is an absolute minimum. From well-known least squares theory we know that $\hat{\xi}$ satisfies the consistent linear system

$$A'(\eta - A\hat{\xi}) = 0, \quad (2)$$

which for the case $n = m$ is equivalent to

$$\eta - A\hat{\xi} = 0.$$

An iterative method to determine $\hat{\xi}$ is presented. The basis for the method is a systematic reduction of the sum of squares of residuals by which trial vectors for $\hat{\xi}$ may be improved to a desired accuracy.

Development of the Method

Suppose that the iteration has proceeded to the p^{th} trial vector, ξ_p and the residual vector now is $\rho_p = \eta - A\xi_p$. In the $(p+1)^{\text{th}}$ trial vector we wish to correct further m_1 components of ξ_p , all the others remaining uncorrected. Say these components of ξ_p are the t_q^{th} components, $q = 1(1)m_1$. A particular way of carrying out this correction is the following

Let

$$\xi_{p+1} = \xi_p + \Delta_{p+1} \quad (3)$$

where the elements of the column vector $\Delta_{p+1} = (d_j)$, $j = 1(1)m$, are to be determined for $j = t_q$, $q = 1(1)m_1$, all the other d_j being set equal to zero; then, by letting $\xi = \xi_{p+1}$ in (1), we have

$$A \Delta_{p+1} + \rho_{p+1} = \rho_p, \quad (4)$$

where the column vector ρ_{p+1} is as yet undetermined. We now form the column vector $\delta_{p+1} = (d_{tq})$, $q = 1(1)m_1$, from the undetermined elements of Δ_{p+1} . We also let α_j , $j = 1(1)m$, denote the column vectors of A , and let A_{p+1} be the matrix (α_{t_q}) , $q = 1(1)m_1$. It is then seen that

$$A \Delta_{p+1} = A_{p+1} \delta_{p+1}. \quad (5)$$

Hence (4) may be stated

$$A_{p+1} \delta_{p+1} + \rho_{p+1} = \rho_p. \quad (6)$$

We now determine δ_{p+1} such that $\rho_{p+1}' \rho_{p+1}$ is a minimum; thus, as in (2)

$$A_{p+1}' (\rho_p - A_{p+1} \delta_{p+1}) = 0, \quad (7)$$

and hence

$$\delta_{p+1} = (A'_{p+1} A_{p+1})^{-1} A'_{p+1} \rho_p. \quad (8)$$

We might add that since $A'A$ is positive definite, $A'_{p+1} A_{p+1}$ is also positive definite.

We now develop a recursive relation for the sum of squares of residuals. From (6) we have

$$\rho'_{p+1} \rho_{p+1} = (\rho_p - A_{p+1} \delta_{p+1})' (\rho_p - A_{p+1} \delta_{p+1}). \quad (9)$$

Substituting from (7) for the indicated product in (9), we have

$$\rho'_{p+1} \rho_{p+1} = \rho'_p \rho_p - \delta'_{p+1} A'_{p+1} \rho_p. \quad (10)$$

With further substitution from (7) we have

$$\rho'_{p+1} \rho_{p+1} = \rho'_p \rho_p - \delta'_{p+1} A'_{p+1} A_{p+1} \delta_{p+1}. \quad (11)$$

From (5) and (11) we then have

$$\rho'_{p+1} \rho_{p+1} = \rho'_p \rho_p - \Delta'_{p+1} A'A \Delta_{p+1}. \quad (12)$$

which is the desired recursive relation.

Since $A'A$ is positive definite, it is at once seen from (12) that $\rho'_{p+1} \rho_{p+1}$ is never greater than $\rho'_p \rho_p$. The sequence $\rho'_p \rho_p$, $p=1,2,\dots$ is therefore monotonic decreasing in the wider sense, and since zero is obviously a lower bound, the sequence has a finite limit. The iterative process is now defined in such a manner that

$$\lim_{p \rightarrow \infty} \rho'_p \rho_p = (\eta - A\hat{\xi})' (\eta - A\hat{\xi}), \quad (13)$$

where $\hat{\xi}$ is the desired solution of (2).

Denote the sequence of iterations by $\dot{z}_p$, $p = 1, 2, \dots$. We agree that the iterations are to be carried out in the following manner. We can always choose a fixed finite number w of iterations, $\dot{z}_p$, $p = r(1) r+w$, for any r , such that every column of A is included at least once in the columns of the matrices A_p , $p = r(1) r+w$. We now show that this sequence of iterations leads to (13).

From (6) and (7) we have

$$A'_u \rho_u = 0,$$

and hence

$$A'_u \rho_v = A'_u (\rho_v - \rho_u). \quad (14)$$

Now let $u = v+1(1) v+1+w$; the w iterations thus considered include every column of A in the columns of the matrices A_u , $u = v+1(1) v+1+w$. It then follows from (14) that

$$\alpha'_j \rho_v = \alpha'_j (\rho_v - \rho_{v+1+f}), \quad \begin{aligned} 0 \leq f \leq w, \\ j = 1(1)m, \end{aligned} \quad (15)$$

where the α_j , $j = 1(1)m$, denote the columns of A . Hence, from (15)

$$A' \rho_v = \|\alpha'_j (\rho_v - \rho_{v+1+f})\|, \quad \begin{aligned} 0 \leq f \leq w, \\ j = 1(1)m. \end{aligned} \quad (16)$$

Now, it may be seen from (4) that

$$\rho_v - \rho_{v+1+f} = A \sum_{r=1}^{1+f} \Delta_{v+r}, \quad f \geq 0,$$

and further, since by definition $\rho_v = \eta - A\xi_v$, (16) may be stated

$$A' (\eta - A\xi_v) = \|\alpha'_j A \sum_{r=1}^{1+f} \Delta_{v+r}\|, \quad \begin{aligned} 0 \leq f \leq w, \\ j = 1(1)m. \end{aligned} \quad (17)$$

Thus,

$$\xi_v = (A'A)^{-1} A'\eta - (A'A)^{-1} \left\| \alpha'_j A \sum_{r=1}^{l+f} \Delta_{v+r} \right\|, \quad \begin{matrix} 0 \leq f \leq w, \\ j = 1(1)m. \end{matrix} \quad (18)$$

Now, since the sequence $\rho'_u \rho_u$, $u = 1, 2, \dots$ has a limit, it follows from (12) that $\lim_{u \rightarrow \infty} \Delta'_u A'A \Delta_u = 0$. Furthermore, since $A'A$ is positive definite, it necessarily follows that $\lim_{u \rightarrow \infty} \Delta_u = 0$. We then have that

$$\lim_{v \rightarrow \infty} \sum_{r=1}^{l+f} \Delta_{v+r} = 0, \quad 0 \leq f \leq w. \quad (19)$$

It is now seen from (18) and (19) that

$$\lim_{v \rightarrow \infty} \xi_v = (A'A)^{-1} A'\eta. \quad (20)$$

Comparing (20) and (2), it is seen that $\lim_{v \rightarrow \infty} \xi_v$ is the desired solution.

We have stated that the sequence $\rho'_p \rho_p$, $p = 1, 2, \dots$ is monotonic decreasing in the wider sense. In practice, we want this sequence to be monotonic decreasing since then there is improvement in the trial vector at each iteration. This is simply done by choosing A_{p+1} different from A_p .

In conclusion, an iterative method for solving systems of linear equations has been outlined. The method assures improvement in trial solutions at each iteration.

Remarks on Application of the Method

In hand computations, where one has relative freedom in guiding the computations, it is suggested that the method be applied as follows:

We let each A_p be a single column of A . Thus, let A_{p+1} be α_s , the s^{th} column of A . δ_{p+1} is now the correction to be added to x_s , the s^{th} component of the current trial vector ξ_p . From (8) we then have

$$\delta_{p+1} = \alpha'_s \rho_p / \alpha'_s \alpha_s. \quad (21)$$

From (8) and (11), remembering that δ_{p+1} is now a scalar, we have that the $(p+1)^{\text{th}}$ sum of squares of residuals is given by.

$$\rho'_{p+1} \rho_{p+1} = \rho'_p \rho_p - (\alpha'_s \rho_p)^2 / \alpha'_s \alpha_s. \quad (22)$$

Now (21) and (22) immediately suggest that it is advantageous to normalize initially the column vectors of A so that $\alpha'_s \alpha_s = 1$. In this manner many divisions may be eliminated in the iterations.

Once having normalized the columns of A , it is seen from (22) that an advantageous choice for a column of A is that column which has the largest absolute scalar product with the current residual vector. It would appear in some cases that an inspection of the numbers involved may indicate the more favorable choices; to say the least, such an inspection may indicate which column vectors should not be chosen for the next iteration.

We now indicate another application of the method. This application is best illustrated by an example; the general case is easily inferred.

We partition A as follows:

$$A = \left(\begin{array}{c|c|c} A_1 & A_2 & A_3 \end{array} \right).$$

We use A_1 , A_2 , and A_3 in the first, second, and third iterations in that order. This completes a cycle which gives corrections to be added to each element of the p^{th} trial vector ξ_p . Denote δ_1 , δ_2 , δ_3 to be the column vectors of corrections obtained from the first, second, and third iterations, respectively.

It is seen from the iterations that we have the following equations for a cycle:

$$\begin{aligned} A_1' A_1 \delta_1 &= A_1' \rho_p, \\ A_2' A_2 \delta_2 &= A_2' (\rho_p - A_1 \delta_1), \\ A_3' A_3 \delta_3 &= A_3' (\rho_p - A_1 \delta_1 - A_2 \delta_2). \end{aligned} \quad (23)$$

Let us define $B_{uv} = A_u A_v$, $u, v = 1, 2, 3$.

With this nomenclature, by simple operations (23) is reduced to

$$\begin{pmatrix} I & 0 & 0 \\ B_{22}^{-1} B_{21} & I & 0 \\ B_{33}^{-1} B_{31} & B_{33}^{-1} B_{32} & I \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix} = \begin{pmatrix} B_{11}^{-1} A_1 \\ B_{22}^{-1} A_2 \\ B_{33}^{-1} A_3 \end{pmatrix} \rho_p. \quad (24)$$

Thus, (24) describes the solution of a complete cycle. It is to be noted that the square matrix on the left is triangular and hence easily inverted, particularly if the unit partitions on the diagonal are not many in number. With this inversion, (24) is reduced to

$$\delta_{p+1} = C \rho_p, \quad (25)$$

and if so desired, matrix-squaring methods may be used to accelerate the iteration.

It is to be noted from (24) that if A_u , $u = 1(1)m$, are the single columns of A , this method of iteration is identical to the Gauss-Seidel iteration applied to the system $A'A\xi = A'\eta$.

It is also worthwhile to notice that letting A_u , $u = 1(1)q$, contain more than one column of A must of necessity increase the convergence rate of the iteration at the expense of added initial work in setting up the cycle calculations in (24). This amounts to increasing the degrees of freedom of the minimization of the sum of squares of residuals at each iteration. It would appear that in applying the method to matrix inversion where in effect the same matrix multipliers in (24) are used to arrive at corrections for each column of the trial inverse, the increase in convergence rate by the means described is of interest.

Reference

For a geometrical discussion of the presented iterative method for solving linear systems, reference is made to Some Numerical Method for Solving Systems of Linear Equations, A. S. Householder, the American Mathematical Monthly, Vol. LVII, No. 7, August-September, 1950.

END OF DOCUMENT