IV.—Studies in Practical Mathematics. V. On the Iterative Solution of a System of Linear Equations.* By

A. C. Aitken, D.Sc., F.R.S., Mathematical Institute, 16 Chambers Street, Edinburgh, 1.

(MS. received April 28, 1949. Read July 4, 1949)

Synopsis

The convergence of customary processes of iteration for solving linear equations, in particular simple and Seidelian iteration, is studied from the standpoint of matrices. A new variant of Seidelian iteration is introduced. In the positive definite case it always converges, the characteristic roots of its operator being real and positive and less than unity.

1. Preliminary Considerations

There has appeared recently a series of papers (Bodewig, 1947) reviewing various extant methods of solving simultaneous linear equations and assessing their relative efficiency. It is claimed, on the basis of this assessment, that the oldest and most elementary of these methods, that of successive elimination of the unknowns followed by resubstitution, involves fewer and simpler operations than any of its more recent competitors. The assessment of efficiency is based on the number of operations required before the solution is complete; for example, it is stated that in the case of a system of n equations, the whole process of solution by the method of elimination requires $n(n^2+3n-1)/3$ multiplications and $n(n-1)(2n+5)/6$ additions.

These assessments must be viewed with respect; but are equally subject to qualification and revision. For example, many operations of addition are not separate from those of multiplication, but on the machine are cumulated along with them. Again, copying down is itself an operation that takes a relatively large proportion of time and involves a risk of error; any operation that minimizes the necessity of such copying is advantageous. There are certainly cases, especially when the matrix A of the system is dominated by its diagonal elements, in which Seidelian iteration converges well. Further, if in such a case the greatest characteristic root of the iterative operation is real and rather small compared with 1, powerful methods are available for gaining at one step a much enhanced approximation. Finally, the Southwellian

* This paper was assisted by a grant from the Carnegie Trust for the Universities of Scotland.

technique, which is that of Seidelian iteration with an admixture of opportunism, has stood the test of trial in the engineering applications with which it deals.

In the present paper, we first make some general observations on the rapidity of convergence of iterative processes (some of this crosses ground already traversed by Dr Bodewig), and then go on to introduce a new variant of the Seidelian process, based on an operator which, when the matrix A of the system to be solved is positive definite, has characteristic roots real and confined to the range $0 \leqslant \lambda < 1$. The advantages of this will be seen in due course.

For the sake of economy of space, the numerical examples refer to systems of low order, and are indeed capable of being solved just as rapidly in other ways. They are intended to serve as illustrations of principle, and are designed to exhibit to the eye, in a not specially favourable case, those properties of convergence that are derived in the text.

2. The General Process of Iteration

The usual processes of iteration have often been described (*e.g.* in Frazer, Duncan and Collar, 1938; Bodewig, 1947) and are easily classified. Let the system of equations be denoted in matrix notation by $Ax=h$. The matrix A can be expressed in infinitely many ways as $B-C$. We choose B as a non-singular matrix such that either B^{-1} is easily evaluated as a first step, or else the effect of B^{-1} on any vector can be obtained by simple arithmetical routine. The choice of B is wide; in the simplest case it could be scalar; in the next simplest, diagonal; in the next simplest, upper or lower triangular. Iteration consists in the use of the recurrence relation

$$Bx^{(t+1}=Cx^{(t}+h, \quad \text{that is,} \quad x^{(t+1}=B^{-1}Cx^{(t}+B^{-1}h, \tag{1}$$

the best working rule consisting in the use of vector-differences, thus,

$$B\{x^{(t+1}-x^{(t)}\}=C\{x^{(t}-x^{(t-1}\}. \tag{2}$$

The iterated vectors $x^{(t}$ are derived in this way from an initial vector $x^{(0}$, adopted as a first approximation to the vector x of solutions.

The rapidity of convergence of the sequence $x^{(t}$ towards x thus depends on the latent roots of the matrix $B^{-1}C$. The characteristic equation of the iteration is therefore $|\lambda B-C|=0$, and if its roots, in descending order of moduli, are $\lambda_1, \lambda_2, \ldots, \lambda_n$, convergence will be assured provided that $|\lambda_1|<1$, and will be the more rapid the smaller $|\lambda_1|$ is. In certain cases λ_1 will be real; if this can be ensured, methods for accelerating the convergence can be applied. These methods can, in fact, be applied even when $|\lambda_1| > 1$, provided that $|\lambda_2| < 1$, in much the same way that $1+\lambda+\lambda^2+ \ldots +\lambda^{t-1}$, when provided with the remainder term $\lambda^t/(1-\lambda)$, will yield $1/(1-\lambda)$ for *all* values of λ; but this case has little practical value. If λ_1 is one of a pair of conjugate complex roots, accelerative methods exist (Aitken, 1925, p. 302), but are less convenient to apply.

A first classification of iterative methods may be based on the nature of B. If B is purely diagonal, for example if it is the "vertebra" of A, so that $b_{ii}=a_{ii}$, we have the type of iteration that is often called *simple* or *ordinary*, but could equally well be called *diagonal*. If B is "lower triangular", that is, if $b_{ij}=0$, $i<j$, $b_{ii}\neq 0$, we have what may be called *lower triangular* iteration. Not essentially different is *upper triangular* iteration. Seidelian iteration (Seidel, 1874), as usually understood, is the special lower triangular iteration where $b_{ii}=a_{ii}$, $b_{ij}=a_{ij}$, $i>j$, $b_{ij}=0$, $i<j$; or any similar iteration with the equations in some permuted order. A variant due to Morris (Frazer, Duncan and Collar, 1938, p. 132) uses a preliminary evaluation of B^{-1} and $B^{-1}h$, but the results at each step are the Seidelian ones.

At this stage we may illustrate, by simple examples of the 3rd order, the characteristic equations of the simplest diagonal iteration $(b_{ii}=a_{ii})$ and of lower triangular Seidelian iteration respectively:

$$\begin{vmatrix} \lambda a_{11} & a_{12} & a_{13} \\ a_{21} & \lambda a_{22} & a_{23} \\ a_{31} & a_{32} & \lambda a_{33} \end{vmatrix} = 0, \qquad (3) \qquad\qquad \begin{vmatrix} \lambda a_{11} & a_{12} & a_{13} \\ \lambda a_{21} & \lambda a_{22} & a_{23} \\ \lambda a_{31} & \lambda a_{32} & \lambda a_{33} \end{vmatrix} = 0. \qquad (4)$$

If A is symmetric, the roots of (3) are real, though not necessarily all less than 1; in (4), on the other hand, where we note the unsymmetrical disposition of the λ's, the roots may be real or complex according to circumstance, even when A is symmetric. In justification of this last remark we may refer to the simple positive definite example of the 3rd order, $a_{ii}=3$, $a_{ij}=a_{ji}=2$, $i\neq j$.

In simple iteration it is known (Bodewig, 1947), and will be proved here, that if A is symmetric and *positive definite*, it is always possible to find a diagonal matrix B such that the numerically greatest latent root λ_1 of the iterating operator is not only real but such that $|\lambda_1| < 1$. It is also known (Whittaker and Robinson, 1929, p. 255; Bodewig, 1947) that in this positive definite case Seidelian iteration always converges. We proceed to prove these facts from the standpoint of matrix theory.

3. The Positive Definite Case

Let A be positive definite. Then its diagonal elements a_{ii} are all positive and constitute a diagonal matrix D. We may normalize the given equations $Ax = h$ to

$$D^{-\frac{1}{2}}AD^{-\frac{1}{2}}y = D^{-\frac{1}{2}}h, \quad \text{where} \quad y = D^{\frac{1}{2}}x. \tag{1}$$

This normalization is for the purpose of theory only, and need not be resorted to in practice. There, it is enough to semi-normalize, obtaining the non-symmetric system $D^{-1}Ax = D^{-1}h$. However, since

$$D^{-1}A = D^{-\frac{1}{2}}(D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}}, \tag{2}$$

the matrices $D^{-1}A$ and $D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ are similar and so have the same latent roots. Each has unit diagonal elements, and so the trace is equal to n; hence, the latent roots being all positive, we must have $\lambda_1 < n$, $\lambda_n > 0$. By writing $D^{-1}A = \frac{1}{2}n(I - C)$, $h = \frac{1}{2}nk$, we reduce the equations to the form $(I - C)x = k$, where the latent roots of C evidently lie in the range $-1 < \lambda < 1$. A simple iteration based upon

$$x^{(t+1)} = Cx^{(t)} + k \tag{3}$$

can now be applied. The process will always converge; though often, as we shall see (§ 6) by an example, with disappointing slowness. Naturally, if we knew beforehand the approximate location of the latent roots of $D^{-1}A$, and especially of the smallest root, we could in most cases make a better change of matrix origin than $\frac{1}{2}nI$; but such knowledge is not usually precise enough.

Let us next consider the Seidelian case. The property of convergence, in the case when A is positive definite, is perhaps most easily deduced indirectly, from the fact that $Ax = h$ can be regarded, and in infinitely many ways, as the normal equations of some linear least-square problem. For we have $A = M'M$, $|M| \neq 0$, a resolution which is possible in infinitely many ways, since HM, where H is an arbitrary orthogonal matrix, can replace M here; and then the equations $M'Mx = h$ are the normal equations corresponding to the "observational" equations $Mx = (M')^{-1}h$.

These normal equations arise from the minimizing of the definite quadratic form $s^2 = (Mx - k)'(Mx - k)$, where $k = (M')^{-1}h$, and this is a sum of squares which can be transformed to a different sum of squares (Whittaker and Robinson, 1929, p. 255) by the classical reduction of Lagrange (Turnbull and Aitken, 1932, p. 83), in which, if we gather all terms in x_1 into a squared term, we obtain

$$s^2 = a_{ii}^{-1}(a_{11}x_1 + a_{12}x_2 + \ . \ . \ . \ + a_{1n}x_n - h_1)^2 + \ . \ . \ ., \tag{4}$$

every term later than the first being free of x_1. To annul the squared residual bracketed in the first term is to *reduce* s^2; and such annulling is the typical single operation in any phase of Seidelian iteration. The point is, that the annulment is effected by modifying x_1 only, and so the later terms in s^2 are unaffected. Each operation of Seidelian iteration thus reduces s^2; and so the vector of values $\{x_1 x_2 \ldots x_n\}^{(t}$ converges, as desired, to those values that minimize s^2. It follows, *a posteriori*, that the latent roots λ_i of the Seidelian operation must be such that $|\lambda_i| < 1$; but, as has been seen from the simple example of § 2, they may be complex.

The Seidelian operation can be described thus. Take $d_{ii} = a_{ii}$ as before. Let C be the lower triangular matrix $c_{ii} = 0$, $c_{ij} = a_{ij}$, $i > j$. Then $A = D - C - C'$, the iterating matrix is $(D-C)^{-1}C'$, and the characteristic equation of the iteration is $|\lambda(D-C) - C'| = 0$. We learn therefore that if $D - C - C'$ is positive definite, then the roots of the above equation are such that $|\lambda| < 1$.

4. The Acceleration of Convergence

The successive vector increments are derived, as we have seen in § 2, by the recurrence

$$x^{(t+1} - x^{(t} = B^{-1}C\{x^{(t} - x^{(t-1}\}. \tag{1}$$

We are therefore on the familiar and well-explored ground of the repeated matrix operation on a vector, and we know that in the case that is being considered

$$x^{(t+1} - x^{(t} = \lambda_1\{x^{(t} - x^{(t-1}\} \tag{2}$$

will hold with greater and greater approximation, and the more if $|\lambda_2|/|\lambda_1|$ is small. In practice it is advantageous to form the vector-difference at the earliest possible stage and to use it as operand, thus dismissing h from consideration. We shall therefore use

$$B\Delta x^{(t} = C\Delta x^{(t-1}, \quad \text{where} \quad \Delta x^{(t} = x^{(t+1} - x^{(t}, \tag{3}$$

as the recurrence. When the iterated differences are adequately small, we cumulate them upon $x^{(0}$; and the errors of rounding-off likely to be incurred in such a cumulation can be obviated by retaining one or two additional digits, the customary expedient for "stabilizing" any calculation.

In this positive definite case, however, provided that λ_1 is real, it will usually not be necessary to continue the iteration until the vector-differences are as small as desired; it will be enough that the corresponding elements in consecutive vector-differences should begin to show an approach to a geometrical progression, which will be of common ratio λ_1. The accelerative methods are then available. Suppose in fact that $\Delta x_i^{(t-1}$ and $\Delta x_i^{(t}$ are two such corresponding elements; write $\Delta^2 x_i^{(t-1} = \Delta x_i^{(t} - \Delta x_i^{(t-1}$, and form the quotient

$$\{\Delta x_i^{(t)}\}^2 / \Delta^2 x^{(t-1}. \qquad (4)$$

This (Aitken, 1925, p. 301; Holme, 1932; Steffensen, 1933) will approximate to the remainder term; it should be applied, therefore, to each x_i in the vector of solutions derived by cumulating up to $x^{(t}$, and the resulting vector can then be tested by further Seidelian iteration. A somewhat less accurate remainder term is given by the quotient

$$\{\Delta x_i^{(t-1} \Delta x_i^{(t)}\} / \Delta^2 x^{(t-2}, \qquad (5)$$

and this will serve sometimes as a check.

5. A Modification of Seidelian Iteration

It has been established that in the symmetric and positive definite case Seidelian iteration converges. If a type of Seidelian iteration could be devised such that all the roots of the characteristic equation were *real*, the accelerative methods would in every case be available, and would enhance the already existing advantage of convergence. We suggest therefore the following procedure.

As before, we have $A = D - C - C'$, positive definite. Let us begin Seidelian iteration as usual, obtaining from an initial vector $x^{(0}$ the improved values of the unknowns, $x_1^{(1}, x_2^{(1}, \ldots, x_n^{(1}$. Then, instead of beginning (as is the usual procedure) a new cycle $x_1^{(2}, x_2^{(2}, \ldots$, let us go *back through the unknowns in reverse order*, $x_{n-1}^{(2}, x_{n-2}^{(2}, \ldots, x_1^{(2} = x_1^{(3}$; then down again, $x_2^{(3}, x_3^{(3}, \ldots, x_n^{(3} = x_n^{(4}$; then up again, and so on. This is a to-and-fro or two-phase Seidelian iteration, and the results are obtained in the order indicated below: *e.g.*

$$\begin{array}{c|ccccccccccc|}
 & & x_1^{(1} & & x_1^{(2} & = & x_1^{(3} & & x_1^{(4} & = & x_1^{(5} \\
x_2^{(0} & \downarrow & x_2^{(1} & \uparrow & x_2^{(2} & \downarrow & x_2^{(3} & \uparrow & x_2^{(4} & \downarrow & \\
x_3^{(0} & & x_3^{(1} & & x_3^{(2} & & x_3^{(3} & & x_3^{(4} & & \\
x_4^{(0} & & x_4^{(1} & = & x_4^{(2} & & x_4^{(3} & = & x_4^{(4} & &
\end{array} \qquad (1)$$

The two stages, down and up, are to be regarded as the complementary halves of a double Seidelian operation, rather like the two half-oscillations of a complete oscillation; so that it will be useful to speak of elements

$x_i^{(t}$, $x_i^{(t+2}$, $x_i^{(t+4}$, . . . as being "in phase". Forming therefore the differences of elements in phase, let us say

$$\Delta x_i^{(t} = x_i^{(t+2} - x_i^{(t}, \tag{2}$$

we can regard the corresponding vector-differences as operands, and the complete operation is then characterized by the matrix

$$(D - C')^{-1}C.(D - C)^{-1}C'. \tag{3}$$

It will now be shown that every latent root of this matrix is real and such that $0 \leqslant \lambda < 1$. For in the first place $|\lambda| < 1$, since each individual operation is of Seidelian type. Next, we have

$$(D - C')^{-1}C.(D - C)^{-1}C' = D^{-\frac{1}{2}}\{(I - K')^{-1}K(I - K)^{-1}K'\}D^{\frac{1}{2}}, \tag{4}$$

where $K = D^{-\frac{1}{2}}CD^{-\frac{1}{2}}$. This again, since K and $(I - K)^{-1}$ are permutable, is similar to $(I - K)^{-1}KK'(I - K')^{-1}$, namely a non-negative definite matrix of form $M'M$ and in general of rank $n - 1$, since K is in general of rank $n - 1$. We conclude that $0 \leqslant \lambda < 1$. The effective positiveness of the roots confers some arithmetical advantage.

It may be of interest to derive the above result from first principles. To each of the latent roots λ there corresponds a non-trivial characteristic vector q such that

$$(I - K')^{-1}(I - K)^{-1}KK'q = \lambda q. \tag{5}$$

Hence
$$q'KK'q = \lambda q'(I - K)(I - K')q,$$

that is,
$$\lambda = q'KK'q/\{q'(I - K - K')q + q'KK'q\}. \tag{6}$$

The quadratic form $q'(I - K - K')q$ is positive definite, by hypothesis, and $q'KK'q$ is non-negative definite. It follows that $0 \leqslant \lambda < 1$.

In the practical technique we begin with $x^{(0}$, form $x^{(1}$ with the downward operations and then $x^{(2}$ with the upward ones. Constructing then the vector-difference $x^{(2} - x^{(0}$, we take it as operand for the further downward and upward operations. When convergence seems adequate, we can cumulate in two ways, first upon $x^{(0}$ with all vector differences that are in phase with it, and again upon $x^{(1}$ with all vector-differences in phase with it; these two vectors of iterated solutions can be used to check each other.

6. Numerical Examples of the Various Iterations

To solve

$$\begin{bmatrix} 3{\cdot}17 & 0{\cdot}92 & -1{\cdot}07 & 1{\cdot}13 \\ 0{\cdot}92 & 3{\cdot}86 & -0{\cdot}89 & -0{\cdot}77 \\ -1{\cdot}07 & -0{\cdot}89 & 5{\cdot}14 & 1{\cdot}79 \\ 1{\cdot}13 & -0{\cdot}77 & 1{\cdot}79 & 6{\cdot}23 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 8{\cdot}08 \\ 6{\cdot}32 \\ 5{\cdot}58 \\ 11{\cdot}05 \end{bmatrix},$$

beginning with values 1·5, 1·5, 1·0 as first approximations for x_2, x_3, x_4. The usual Seidelian iteration gives the first iterated vector {2·2634 1·6432 1·4931 1·1372}. Forming the first vector-difference, we continue:

							Check
	−928	−479	−151	−52	−27	2·0997	8·0796
1432	479	58	15	4	1	1·6989	6·3200
− 69	−588	−228	− 85	−29	−15	1·3986	5·5797
1372	396	160	54	18	9	1·2009	11·0496

Stopping at the stage where each entry seems to be about one-third of the corresponding entry in the column before, we apply the corrections of § 4 (4), for example $-52^2/99=-27$, and cumulate upon the earliest values 2·2634, 1·5, 1·5 and 1·0. The check suggests that the solutions have at most small errors in the 4th decimal place. The accurate values are in fact {2·0999 1·6989 1·3987 1·2009}.

The new variant of the Seidelian iteration gives the following opening values, down and then up, after which we continue with vector-differences in phase:

	2·2634			−1132		−347		−108	−49	2·0998
1·5	1·6432	1·6579	1579	417	285	98	86		37	1·6987
1·5	1·4931	1·4453	− 547	−643	−317	−209	−102		−48	1·3986
1·0		1·1372	1372		442		135		59	1·2008

For example, $1{\cdot}5+0{\cdot}1579+0{\cdot}0285+0{\cdot}0086+0{\cdot}0037=1{\cdot}6987$.

Here there is little to choose between the old and the new Seidelian iteration, for it so happens that the largest characteristic root of the former is real and fairly small.

Simple iteration, on the other hand, when performed with the change of matrix origin mentioned in § 3, so that for example the first equation of iteration is

$$6{\cdot}34x_1^{(1}=3{\cdot}17x_1^{(0}-0{\cdot}92x_2^{(0}+1{\cdot}07x_3^{(0}-1{\cdot}13x_4^{(0},$$

converges, but so much more slowly than the Seidelian iterations that it seems unnecessary to give the details. In particular, if we begin with the initial values {2·0 1·5 1·5 1·0}, the successive increments for x_1 take a long time in settling down to an approximate geometrical progression, of common ratio about 0·75.

REFERENCES TO LITERATURE

AITKEN, A. C., 1925. "On Bernoulli's Numerical Solution of Algebraic Equations", *Proc. Roy. Soc. Edin.*, XLVI, 289–305.

——, 1937. "The Evaluation of the Latent Roots and Latent Vectors of a Matrix", *Proc. Roy. Soc. Edin.*, LVII, 269–304.

BODEWIG, E., 1947. "Bericht über die verschieden Methoden zur Lösung eines System linearer Gleichungen mit reellen Koeffizienten", *Indag. Math.*, IX, 441–452, 1104–1116, 1285–1295; X, 53–64, 211–219.

FRAZER, R. A., DUNCAN, W. J., and COLLAR, A. R., 1938. *Elementary Matrices*, Cambridge University Press.

STEFFENSEN, J. F., 1933. "Remarks on Iteration", *Skandinavisk Aktuarietidskr.*, 64–72.

TURNBULL, H. W., and AITKEN, A. C., 1932. *Theory of Canonical Matrices*, Blackie & Son, London and Glasgow.

WHITTAKER, E. T., and ROBINSON, G., 1924. *The Calculus of Observations*, Blackie & Son, London and Glasgow.

(*Issued separately June* 16, 1950)