PRACTICAL PROBLEMS OF MATRIX CALCULATION*

HAROLD HOTELLING

UNIVERSITY OF NORTH CAROLINA

1. Introduction

Statistical analysis involving any considerable number of variates leads usually to calculations which, if not well organized and conducted with the help of suitable machines, may be excessively laborious. The great possibilities opened up by the advances in the theory of multivariate statistical analysis will gain in accessibility as computational methods improve in this domain. The computations most needed involve matrices. This is true, for example, in the method of least squares and in calculating multiple correlation coefficients; it is true in the calculation of the generalized Student ratio and figurative distance that has become the modern substitute for Karl Pearson's coefficient of racial likeness; also in studying the relations between two or more sets of variates, and the principal components of one set.

The same computational problems arise also in many fields outside of statistics—if indeed we can speak of any field as being outside of statistics! Thus the study of vibrations in airplanes and other machines and structures; the analysis of stresses and strains in bridges, of electrical networks, of mass spectroscopy of petroleum products; and many other subjects of importance require calculations of the same kinds as multivariate analysis in statistics.

The calculations principally needed are of three main types:

a) Multiplication of matrices, that is, formation of sums of products.

b) Inversion of matrices and solution of systems of linear equations.

c) Determination of latent roots and latent vectors, also known as characteristic roots and vectors, as principal components, and by such mongrel terms as eigenvectors and eigenvalues. More generally, determination of values of $x_1, \cdots, x_p, \lambda$ satisfying

$$\sum_j (a_{ij} - \lambda b_{ij})x_j = 0\,, \qquad i, j = 1, \cdots, p.$$

In addition there are special matrix computational problems such as those encountered by Koopmans [8].[1] It might be thought that the calculation of determinants should have an important place here. However, in most problems of practical importance of which I am aware, if a determinant occurs in a formula it is associated with other determinants in such a way that all that is really needed is one or more of the ratios of determinants forming the inverse of a matrix. Hence these problems come under (*b*). For example, one of the

* A revision of a Symposium lecture, taking account of developments in the subject down to May, 1946, and, with an added note, to June, 1948.

[1] Boldface numbers in brackets refer to references at the end of the paper (p. 293).

expressions for the multiple correlation coefficient involves the ratio of two determinants, of which one is a principal minor of the other. This ratio, or its reciprocal, is one of the elements of the inverse of the matrix of the larger determinant, which is easier to compute than the two separate determinants. The reason is essentially the repetition of labor involved in computing separately two determinants many of whose elements are common. Furthermore, if the multiple correlation coefficient is wanted, there is generally a need also for several other things that can be found easily from the inverse of this same matrix.

In a paper [5] published in 1943 I went over the subject of matrix calculation in much more detail than is possible in a single lecture. I shall try not to repeat what I then wrote except insofar as is necessary for intelligibility and continuity. Instead, I shall concentrate mainly on developments of which I have learned more recently, while including some slight revision of the former treatment at certain critical points and a little speculation about future developments.

The interesting questions regarding calculation with matrices pertain to those with large numbers of rows and columns. For matrices of two, three, and often four rows there appears to be little or no gain in any methods departing materially from the straightforward use of the definition of the function to be computed. But the labor of inverting a matrix or evaluating a determinant or solving a system of linear equations by direct methods increases approximately as the cube of the number of rows or of unknowns, and when this number is large the advantage of oblique and iterative methods may be very great.

2. Mechanical and electrical devices

Choice of methods is inevitably connected with the machines available or capable of being made available. One fundamental desideratum is a better digital device for multiplying large matrices than any now available. Multiplication of matrices is essentially what is needed to get the covariances and correlations among a set of observed variates. It is also important in connection with our other two fundamental classes (*b*) and (*c*) of problems, since some of the leading methods of dealing with these involve matrix multiplication.

Machines of the Monroe, Marchant, and Fridén types I propose to call "calculators" until a better name is found, to distinguish them on the one hand from adding machines and on the other from more complicated machines, electrical networks, and other kinds of devices. For finding single isolated sums of products or correlation coefficients it does not appear that anything better than a modern calculator is now or is likely soon to become available. These calculators also have an advantage over alternative methods of computing all the sums of squares and of products for two or three variates. Recent Monroe models have for this purpose a distinct advantage over the others in the squaring lock. Unfortunately these modern high-speed machines seem to break down rather frequently, requiring a good deal of repair work.

The disadvantage of calculators of this type when matrices of any size are to be multiplied is that the same rows of numbers must be set in by hand again and again as multipliers, and columns of numbers must be set in likewise as multiplicands without taking advantage of the fact that the same columns are used over and over. In other words, the calculator has too limited a memory for such work. In order to improve its memory, numbers may be stored in the form of holes at suitable places in tapes or cards, or in some other form that can be recalled en masse and read and utilized by the machine again and again. The chance of human error is then confined to the correct copying of each number only once.

The first utilization of number-storage for matrix calculation seems to have been in a machine reported in 1925 by Clark L. Hull [7]. Punched tapes were used in conjunction with a device attached to a calculator. The machine read the numbers by poking steel fingers through the holes. While a sum of products was being formed, the operator could go out to lunch, leaving the machine alone in a locked room to stop itself at the conclusion of the operation. Two of these machines were built, one for the University of Wisconsin and one for the National Research Council. The cost of a future machine of this kind was estimated at $1200. However, the idea was abandoned in favor of punched-card machines. A source of trouble seems to have been the tearing of the tapes, though these were of tough kraft paper.

Hollerith cards are capable of carrying as many as ninety variates on a card. With suitable equipment the product of any two of these variates can be formed, summed over a large number of cards at very high speed, and presented simultaneously on dials, in print, and in the form of other card punches. This seems to offer almost an ideal solution of the problem of matrix multiplication except for the cost and ponderous character of the machines, and the fact that after each product-sum is formed it is necessary to do a good deal of resetting of the machine to prepare for the next. There are also other ways of forming product-sums with punched cards, but they all have these difficulties. The next step forward in this direction would appear to be the development of a device for automatically resetting the machine so as to get the product matrix without any human intervention between placing two sets of cards representing the factors in the machine and getting out a printed and punched-card product matrix.

With such a device, which does not seem out of the question in the near future, and which may be actually available in the punched-tape Aiken machine at Harvard, the multiplication of large matrices may become a trifling detail. Mechanical developments of this kind seem capable also of easily forming linear functions of matrices and therefore, in conjunction with the matrix multiplication, polynomials in matrices. But the formation of inverses or principal components is of another order of difficulty, and Hollerith-type machines do not appear to lend themselves readily to direct determinations of these. We must therefore look to other types of mechanical and electrical aids, which, however, may be supplemented by ready matrix multiplication with punched cards or tapes.

A matrix-multiplier designed by L. R. Tucker and built by the International Business Machines Corporation is now in use in the University of Chicago laboratory of L. L. Thurstone. This uses two-digit entries, represented for one factor by wirings of a plug-board and for the other by lengths of soft pencil mark on a sheet of paper. These latter act as resistances when the paper is pressed by a set of contacts, and the product-sum is read to two significant figures on an ammeter. This is a degree of accuracy insufficient for many purposes.

For solving systems of linear equations several electrical devices have been built. One of these is now actually being put on the market by the Consolidated Engineering Company of Pasadena. It is arranged so as to solve as many as twelve linear equations in twelve unknowns, and is essentially an electrical network of resistances alone, in contrast with the earlier Mallock device (reference in [5]), which relies on induction. After setting in the coefficients the operator must try different combinations of settings for twelve knobs, representing the unknowns, until a pointer reaches a zero position indicating that the trial values of the unknowns satisfy the equations. There may be some question whether this process might not require an excessive amount of time with a large number of unknowns, but data supplied by the company regarding the operation of the Pasadena device with certain examples indicate that problems with no more than twelve unknowns can be solved with very satisfactory speed in a few minutes. The coefficients are set in as four-digit numbers, and in the examples given the results are correct to four significant figures except sometimes in the last place. Developing this device and making it generally available represent an important advance.

Various other electrical equation-solvers have been devised but not marketed. The Mallock device previously referred to is a unique specimen at Cambridge, England, for solving ten or fewer equations. It has an iron ring for each unknown, and on the ring ten coils representing the coefficients of this unknown in the ten equations, arranged so that the number of turns actually used in each coil can be set in as the coefficient, with a reversing switch to make possible the use of negative as well as positive numbers. The ten coils representing each equation are connected in series to form a closed and electrically isolated circuit. Alternating current is supplied to an additional coil on one of the rings. The unknowns are proportional to the magnetic flux in the several rings, which may be read off by means of still another coil on each ring, attached to a galvanometer.

At Columbia University Julian Bigelow, of the Division of War Research, has designed an electrical network involving resistances and vacuum tubes only; and F. J. Murray of the Mathematics Department has built a small-scale model of a device in which, as in the Consolidated's computer, trial values are set on knobs and a pointer brought to a minimum position by repeated trials utilizing something like the classical Gauss-Seidel iterative method of solving equations. Both Bigelow's and Murray's computers use only standard parts, which are produced in quantities and can easily be assembled.

In a manuscript and letter dated May 10, 1946, Cecil E. Leith and Quentin A. Kerns, of Oak Ridge, Tennessee, describe an electronic solution-finder for simultaneous linear equations, of which they have built a model for solving five equations in five unknowns.

Numerous mechanical, hydraulic, and electrical devices for solving systems of linear equations have been designed in the past. The older schemes encountered many practical difficulties such as stretching of tapes, leaking of fluids, excessive stresses, strains, and space requirements in systems of jointed rods, and changes in electrical resistances, capacities, and inductances and in magnetic properties with temperature and time.

All these equation-solvers give answers valid only to a limited degree of accuracy. This feature is inherent in the nature of the problem; for even though all the coefficients and constant terms in a set of equations are integers, this may not be true of the solution, which if presented as a decimal may have an infinite number of places of which only the first few can actually be read. Moreover the accuracy of the solution cannot readily be predicted, and the same device may give quite different degrees of accuracy in different problems. Formulas for predicting accuracy involve determinants which are no easier to compute than the solution sought. All this is connected with the fact that the determinant of the given system may be arbitrarily close to zero without its smallness being apparent in advance.

3. Iteration and error control

An alleged solution of a system of linear equations, whether obtained from an electrical network or otherwise, may therefore be regarded with some suspicion until satisfactorily checked. The most obvious check is to substitute the values found in the equations and to inspect the discrepancies between left- and right-hand members. If these discrepancies are all exactly zero the result is certainly satisfactory. But if, as is almost always the case, they are not exactly zero, what can be said about the accuracy of the solution? With some systems of equations very large errors in the unknowns are consistent with very small discrepancies. A check can hardly be called fully satisfactory unless it provides definite upper bounds for the errors in the unknowns. It is further desirable to have a procedure that not only supplies limits to the errors, but when these are too large provides means of cutting them down. Indeed, for any kind of computation it is highly desirable that the routine shall make provision for these two steps:

1. Find a definite and moderate upper bound for the error in each result.
2. Have an iterative method for obtaining greater accuracy when it is needed, while utilizing the approximate results aready reached.

The first of these criteria may in certain cases be varied to require only an upper bound of measurable probability (cf. [11]).

These two criteria should always be kept in view in devising numerical procedures. The first has a certain analogy with the demands of mathematical

statisticians in the last few decades for limits of error expressed, not in terms of unknown "nuisance" parameters or of a priori probabilities, but as definite numbers. In the computational as in the statistical problem we demand a limit for the error in the value we assign an unknown, and we want this limit to be a definite number, not a function of something as uncertain and hard to find as is the principal unknown itself.

Only one method of dealing with systems of linear equations is known to satisfy both the criteria 1 and 2. This method is fundamentally concerned with the inverse of the matrix of coefficients, whose calculation is not strictly necessary for the solution of the equations. However, the practical situation giving rise to the linear equations frequently requires the inverse matrix also, as in the method of least squares and in certain problems in civil and electrical engineering. Since the solution of the linear equations can be found directly and simply from the inverse matrix, the latter should in all such cases be computed first. Any electrical or other method of solving linear equations can be applied to find the inverse of a matrix; for if a_{ij} is the element in the ith row and jth column of the p-rowed square matrix A and c_{ij} the element in like position in $C = A^{-1}$, the unknowns

$$c_{1k}, c_{2k}, \cdots, c_{pk}$$

corresponding to any particular value of k will satisfy the p linear equations

$$\begin{aligned} \Sigma a_{ij}c_{jk} &= 1 \text{ if } i = k, \\ &= 0 \text{ if } i \neq k, \qquad i = 1, 2, \cdots, p. \end{aligned}$$

This means that each column of C may be determined by solving a system of linear equations whose right-hand members consist of zeros and a single 1, and whose left-hand members have as coefficients the elements of the given non-singular matrix A.

The method satisfying our two criteria of definite limits of error and of iterative improvement was set forth in the seventh and eighth sections of the writer's paper [5]. However, there has been an improvement in the limit of error there used as a result of a suggestion by A. T. Lonseth (cf. [6]). The method is as follows:

Starting with an approximation C_0 to the desired inverse of A, calculate

$$D = 1 - AC_0,$$

where 1 stands for the identical or unit matrix of p rows. The labor of calculating D is merely that of substituting the values found for the unknowns c_{ij} back in the equations and finding the discrepancies. The norm of a matrix is defined as the positive square root of the sum of products of the elements by their complex conjugates. For a real matrix this is the same as the square root of the sum of the squares of the elements. The norm of a matrix L is denoted by $N(L)$,

and obviously cannot be exceeded in absolute value by any element of L. Let $N(D) = k$. Then if $k < 1$ it may be proved that

$$N(C_0 - A^{-1}) \leqq N(C_0)k/(1 - k),$$

and this sets a definite upper limit for the greatest error in any element of C_0.

If greater accuracy is needed the following iterative method may be applied. Let

$$C_{m+1} = C_m(2 - AC_m), \qquad m = 0, 1, 2, \cdots.$$

The error in the approximation C_m to A^{-1} is limited by the inequality

$$N(C_m - A^{-1}) \leqq N(C_0)k^{2^m}/(1 - k),$$

provided $k < 1$. The occurrence of the exponential of the exponential on the right means that the number of decimal places of sure accuracy is approximately doubled with each iteration. This geometric increase is in contrast with the behavior of the errors in the better-known classical Gauss-Seidel iteration, where the number of decimal places of accuracy increases in an irregular fashion which on the average is in arithmetic rather than geometric progression, and for which no suitable limits of error are available.

These advantages are gained at the expense, first, of insisting on the use of the inverse matrix, and, second, of a certain risk that the initial approximation may be so poor that the process will not converge at all. In the event of non-convergence, k will come out greater than unity; but also in some cases of convergence k is greater than unity, though similar error norms at later stages become less than unity. When the initial approximation is too poor for convergence, several courses are open for getting a better one. The Gauss-Seidel method may be applied a few times if it seems that the errors are not too great to be corrected in a moderate number of repetitions. A review of the preliminary solution may reveal some gross error whose correction will bring the solution within the necessary limits for convergence. In an extreme case it may be necessary to repeat a straightforward Doolittle solution with an increased number of decimal places. The risk of this most unfortunate contingency is present, however, with any method, since no matter how many decimal places are carried there is always a possibility that the result will show a larger number to have been needed. Another possibility is in partitioning (sec. 4 below).

The probability of convergence in certain cases has been investigated by J. Ullman [**11**]. A proof of the foregoing limits of error is as follows. From the definition of the norm may be obtained fairly easily the two "triangular" inequalities

$$N(A + B) \leqq N(A) + N(B),$$
$$N(AB) \leqq N(A)N(B),$$

where A and B are any matrices, whether square or not, for which the expressions written have meaning in either case. From the last inequality it follows that, for any positive integer m,

$$N(A^m) \leqq [N(A)]^m .$$

Since $N(D)$ was denoted by k, the identity

$$(1 - D)^{-1}D = D + (1 - D)^{-1}D^2$$

shows that

$$N[(1 - D)^{-1}D] \leqq k + kN[(1 - D)^{-1}D] ,$$

whence, if $k < 1$,

$$N[(1 - D)^{-1}D] \leqq k/(1 - k) ,$$

and for any positive integer n,

$$N[(1 - D)^{-1}D^n] \leqq k^n/(1 - k) .$$

Now put

$$D_m = 1 - AC_m , \qquad m = 0, 1, 2, \cdots ,$$

whence

$$C_m = A^{-1}(1 - D_m) .$$

Substitution of this in the equation of iteration,

$$C_{m+1} = C_m(2 - AC_m) ,$$

gives

$$C_{m+1} = A^{-1}(1 - D_m)(1 + D_m) = A^{-1}(1 - D_m^2) .$$

Then

$$D_{m+1} = 1 - AC_{m+1} = D_m^2 ,$$

and since $D_0 = D$ this gives by induction

$$D_m = D^{2^m} .$$

Thus we have for the error matrix,

$$C_m - A^{-1} = -A^{-1}D^{2^m} ,$$

whence

$$N(C_m - A^{-1}) \leqq N(A^{-1})k^{2^m} .$$

Since this limit decreases fairly rapidly when k is materially less than unity, we might be tempted to stop here and to estimate the factor $N(A^{-1})$ as $N(C_m)$. Certainly this would be in keeping with much that is done, but it would violate our first criterion, which requires that the limit of error be *definite*. So long as we are seeking A^{-1} we do not know its norm exactly, and take a chance of

going badly wrong if we replace this by an approximation of uncertain accuracy. In order to remedy the difficulty we recall the definition $D = 1 - AC_0$, whence

$$A^{-1} = C_0(1 - D)^{-1}.$$

Substitution of this in the expression above for the error matrix gives

$$C_m - A^{-1} = -C_0(1 - D)^{-1}D^{2^m}.$$

Applying to this the inequality for the norm of $(1 - D)^{-1}D^n$, we have

$$N(C_m - A^{-1}) \leqq N(C_0)k^{2^m}/(1 - k),$$

which was to be proved.

This limit of error, in addition to being definite and simply expressible in terms of known quantities, and converging rapidly to zero whenever k is appreciably less than unity, has another advantage. It may be used to provide a safe estimate of m, the number of iterations required to attain any required degree of accuracy, as soon as k is known.

In some cases an initial approximation C_0 will not need to be specially calculated but is available from previous work. This is true, for example, if A is a correlation matrix which differs only a little from one whose inverse has previously been found, perhaps because of the inclusion of additional observations or the discovery of errors in previous observations or calculations. Similarly if an electrical or civil engineering design for which the inverse matrix has been calculated is changed slightly, the old inverse supplies a starting-point for iteration to obtain the new one. See also p. 292.

A good matrix-multiplying device will evidently be a great help in iterations of this kind. However, the method has proved its value in numerous problems in which ordinary calculators were the only equipment used. It has been used extensively in the Statistical Research Group devoted to war research at Columbia University, and also by students in my classes since 1936. The orders of the matrices inverted by this method have varied from three to fourteen, though its advantages are usually slight for matrices of fewer than five rows.

The principal equipment appropriate to any place in which any considerable number of matrices are to be inverted or systems of linear equations solved will evidently consist in the future of an electrical equation-solving device together with some means of multiplying matrices. The equation-solver will give quick and easy approximations to the inverse of a matrix, which will then be improved to any desired degree of accuracy, with a check at every stage through the D matrix, by using the matrix-multiplier to substitute in the foregoing iterative formulas.

4. Partitioning

When the order of the matrix to be inverted exceeds the capacity of the electric equation-solver, and in some other cases as suggested above, par-

titioning can be used. If a matrix is partitioned into four submatrices in the form

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

with a and d square and non-singular, and if its inverse is similarly partitioned in the form

$$\begin{bmatrix} A & C \\ B & D \end{bmatrix},$$

then, as remarked by Waugh [12], the inverse may be found by calculating in turn

$$D = (d - ca^{-1}b)^{-1} \qquad B = -Dca^{-1}$$
$$C = -a^{-1}bD \qquad A = a^{-1} - a^{-1}bB\,.$$

If each of these submatrices is of p rows and p columns, use of these formulas means two inversions of p-rowed matrices, together with six multiplications and two subtractions of p-rowed matrices. Since the labor of inversion by direct methods (cf. [13]) varies approximately as the cube of the order of a matrix, the direct inversion of two p-rowed matrices will require only about a quarter of the work of a direct inversion of a single matrix of $2p$ rows. How this saving should be weighed in comparison with the additional multiplications and subtractions and the increased complexity in the arrangement of the work is problematical. The answer must be sought in terms of the equipment available. A good matrix-multiplier is favorable to partitioning, whereas a good electric network of sufficient size to invert the original matrix provides a strong argument against partitioning. But partitioning will be needed when the order of the matrix exceeds the number of equations the network was built to solve directly.

Other methods involving partitioned matrices, including the use of nine instead of four submatrices, have been discussed by W. J. Duncan [2]. A special use of partitioned matrices for solving certain systems of equations, particularly when some but not all of the unknowns are required, was developed by the writer ([5], pp. 22 and 23).

Louis Guttman in a forthcoming paper presents methods of inverting a matrix based on inversion of matrices of successively higher orders. After inverting a two-rowed matrix Guttman performs a series of inversions, using at each stage the inverse already computed of a submatrix of the one currently inverted. He presents three techniques, involving respectively enlargement by a single row and column at a time, enlargement by two rows and two columns at a time, and enlargement by doubling the number of rows and columns.

5. Possible tables of inverse matrices

A table of inverse matrices might possibly be a convenient adjunct, particularly in providing a starting-point for iteration, and for use in connection with partitioning. However, a little consideration shows that a table arranged in the more obvious ways would either be impossibly large or virtually useless because of the wide intervals necessary in the arguments. If a useful table is ever to be produced it must be designed with great finesse.

If each argument is given with a fixed number m of different values, the number of different square matrices of p rows to be listed is m^{p^2}. This is too large, even when $p = 3$, if the arguments range, say, from $-.9$ to $+.9$ by steps of .1. In this case the number of matrices to be taken as arguments is 19^9, or about 322,670,000,000. If twenty of these matrices with their inverses are put on a page, with a thousand pages per volume, this would require about 16,133,500 volumes, enough to fill several very large libraries, and requiring some 500 miles of shelf space.

In order to cut down the number of entries such a table might be limited to correlation matrices of three rows, which are symmetric, positive definite, and have unit elements in the principal diagonal. The inversion of any non-singular matrix A can be reduced by multiplication to that of a positive definite matrix; for

$$A^{-1} = (A'A)^{-1}A',$$

and $A'A$ is positive definite. The unit diagonal elements can be obtained from any positive definite matrix by dividing each element by the geometric mean of the diagonal elements in its row and in its column; this same number is later multiplied into the element in the corresponding place of the inverse of the correlation matrix to obtain the inverse of the original matrix.

The number of independent variables in a three-rowed matrix is now reduced to three. We may omit for the moment the requirement of definiteness and consider a table of all three-rowed correlation matrices. If all three correlations vary independently from $-.9$ to $+.9$ by steps of .1, the number of entries is $19^3 = 6859$. This would still fill a fair-sized volume.

A considerable further reduction in the number of entries will be obtainable by permutation of the rows, and correspondingly of the columns to maintain symmetry, and by changing the signs of some of the rows and of the corresponding columns. All these operations may be applied in reverse to the inverse of the standardized matrix obtained, so as to yield the inverse of the original matrix. By these operations it may be arranged that r_{12} is that one of the correlations having maximum absolute value, and that both r_{12} and r_{13} are positive. Then with the same interval .1 as before, we have the following numbers of matrices to be listed with their inverses:

When $r_{12} = .9$, the possible values of r_{13} are $0, .1, \cdots, .9$, and those of r_{23} are $-.9, -.8, \cdots, +.9$. If these last two correlations are regarded as varying with mutual independence, the total number of entries with $r_{12} = .9$ is $10 \times 19 = 190$. Similarly for $r_{12} = .8$ the total number of matrices to be listed with

their inverses is $9 \times 17 = 153$. For all the values .9, .8, $\cdots$, .1, 0 of r_{12} the aggregate is

$$(1 \times 1) + (2 \times 3) + (3 \times 5) + \cdots + (10 \times 19) = 715 .$$

This is a relatively moderate number, and such a table could easily be put into about 36 pages. Elimination of those matrices which are not positive definite would reduce the size of the table still further.

We must, however, consider whether the values obtained from such a table can be relied upon to provide sufficiently good first approximations to cause the iterative method to converge. For this a sufficient condition (and apparently the only condition easy to apply) is that $N(D) < 1$, where $D = 1 - AC_0$, A is the matrix to be inverted, and C_0 is the initial approximation to its inverse. Now suppose that A_0 is a matrix listed in the table and is used as an approximation to A, so that $C_0 = A_0^{-1}$ is obtained from the table. Then

$$D = (A_0 - A)A_0^{-1} = (A_0 - A)C_0 ,$$

and the process converges if

$$N(A_0 - A)\, N(C_0) < 1 .$$

Since no element of A_0 need differ from the corresponding element of A by more than .05 when the maximum correlation in A does not exceed .95, while the diagonal elements are the same in both cases, we have

$$N(A_0 - A) \leqq \sqrt{6 \times .05^2} = .122 .$$

Hence the process will converge if $N(C_0) < 1/.122 = 8.2$.

This may fail to assure adequate convergence in some cases, since $N(C_0)$ may exceed 8.2. For example, if all the correlations in a three-rowed matrix have the common value r, the inverse is

$$C_0 = \frac{1}{(1-r)(1+2r)} \begin{bmatrix} 1+r & -r & -r \\ -r & 1+r & -r \\ -r & -r & 1+r \end{bmatrix},$$

as may be verified by inspection upon multiplying by the original matrix. The squared norm of C_0 is

$$\frac{3(1+2r+3r^2)}{(1-r)^2\,(1+2r)^2} .$$

This is within the specified limit only if r is either between -1 and $-.5613$ or between $-.4383$ and $+.8225$. The former of these intervals may, however, be

disregarded if the table is to be confined to correlation matrices, since these are positive definite, as our matrix is only when $r > -\frac{1}{2}$. Since the matrix is singular when $r = 1$ or $-\frac{1}{2}$, the iterative method based on a table fails when r is brought too close to either of these values relatively to any fixed tabular interval. But the domain of convergence may be extended considerably if tabular intervals are reduced near the singularities, that is, near the boundaries of the regions within which the determinant is positive. On the other hand, when the correlations are small and the determinant therefore near unity, much broader tabular intervals are appropriate than when the determinant is near zero; and this fact may be used to keep down the size of the table.

In order to illustrate this situation, suppose that

$$A = \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$$

is to be inverted, and that the matrix C_0 written above is taken as a first approximation to A^{-1}. Then

$$D = 1 - AC_0 = \frac{\rho - r}{(1-r)(1+2r)} \begin{bmatrix} 2r & -1 & -1 \\ -1 & 2r & -1 \\ -1 & -1 & 2r \end{bmatrix}.$$

Since the determinant of the matrix last written equals

$$(2r-2)(2r+1)^2,$$

and since its characteristic equation is found by replacing $2r$ by $2r - \lambda$, the latent roots of this matrix are obvious. In order to obtain those of D it is only necessary to multiply them by the scalar multiplier of the matrix last written, obtaining

$$\frac{2(r-\rho)}{1+2r} \quad \text{and} \quad \frac{\rho - r}{1-r},$$

of which the second is a double root. A necessary and sufficient condition for the convergence of the iterative process is known to be that all the latent roots of D be less than unity in absolute value. When $r = 0$ in the present case, this condition is satisfied for $-\frac{1}{2} < \rho < \frac{1}{2}$; and when $r = \frac{1}{2}$ it is satisfied for $0 < \rho < 1$. Thus, so far as positive definite matrices of this particular type are concerned, a table of inverses of matrices need have only two entries, corresponding to $r = 0$ and $r = \frac{1}{2}$, to ensure that an entry can be found from which the iterative process will converge to the true inverse matrix. The test by

means of the norm, which is cruder than that by means of the latent roots though more convenient, guarantees convergence on the basis of such a two-entry table only over the more restricted range of ρ from $-1/\sqrt{6} = -.41$ to $+5/6$.

To devise a useful and feasible table of inverse matrices, if possible at all, is thus largely a matter of suitable selection of the entries in relation to singularities. Provision of a method of convenient interpolation specially adapted to such a table might also help. An adequate and useful table of this kind may not be at all out of the question, in spite of the preposterous size suggested by a first consideration of the problem.

6. Solution of determinantal and associated linear equations

The p homogeneous linear equations

$$\sum_j (a_{ij} - \lambda b_{ij}) x_j = 0, \qquad i = 1, 2, \cdots, p,$$

will possess a solution $x_1, \cdots, x_p$ for each value of the parameter λ satisfying the determinantal equation

$$|A - \lambda B| = 0,$$

where A and B are respectively the matrices of elements a_{ij} and b_{ij}. Direct determination of the p roots λ is laborious, and with the additional work of finding solutions for the linear equations corresponding to each root is very forbidding, when the order p of the matrices is large. However, iterative methods cut down the work substantially, and it will be possible to accelerate their convergence greatly when a satisfactory matrix-multiplying machine is achieved.

The most important, though by no means the only, case occurring in practice is that in which A is symmetric and $B = 1$, so that the latent roots and vectors of A are to be determined. The latent roots of a symmetric matrix are all real, and those of its square are all positive or zero. They may be found iteratively in the following manner. Let X_0 be a column vector (a matrix consisting of a single column of p elements), and let $X_t = AX_{t-1}$ be a sequence of column vectors found by repeated premultiplication by A. Since $X_t = A^t X_0$, this process may be speeded up by raising A to a high power by repeated squaring; this will be particularly advantageous when an efficient matrix-squaring machine becomes available. It is known that the ratio of any element in X_t to the corresponding element in X_{t-1} converges, as t increases, to a latent root λ_1 of A having greatest absolute value. This root will be unambiguous if all the roots of A are positive or zero, as we may assume, for instead of A we may consider A^2. Let us suppose, then, that we are dealing with a symmetric matrix all of whose latent roots are positive or zero.

At the same time that the ratio of corresponding elements of successive vectors is approaching λ_1, the ratios among the elements of a vector are approaching those among values satisfying the corresponding linear equations. These

values are components of a latent vector associated with λ_1. They are not unique, for they may be multiplied by an arbitrary constant. If a constant factor is chosen so as to make the sum of the squares of the elements of the vector unity, it is said to be normalized. The normalized vector is unique except for sign, provided λ_1 is a simple root of the determinantal equation. Let us here for simplicity consider only this case.

After the greatest root λ_1 and the corresponding normalized vector X have been determined with satisfactory accuracy (which remains to be verified), we may form a new matrix A_1 whose latent roots and vectors are the same as those of A excepting that λ_1 is replaced by zero. That these are properties of the matrix

$$A_1 = A - \lambda_1 XX'$$

may be verified as follows. If Y is a latent vector of A different from X, then, as is well known, they are orthogonal, so that $X'Y = 0$. Furthermore, if Y corresponds to the latent root λ_k of A, then $AY = \lambda_k Y$. Consequently, upon postmultiplying by Y the equation defining A_1, we have

$$A_1 Y = AY = \lambda_k Y,$$

so that Y is a latent vector of A_1 associated with the same value λ_k with which it was associated as a latent vector of A. If the equation defining A_1 is postmultiplied by X, we obtain

$$A_1 X = AX - \lambda_1 XX'X.$$

But $AX = \lambda_1 X$; and since X is normalized, $X'X = 1$. Therefore $A_1 X = 0$ verifying that zero is a latent root of A_1, whereas the others are $\lambda_2, \cdots, \lambda_p$.

If now we start with an arbitrary vector and apply the same process of iteration with A_1 as that used previously with A, we shall approach the second greatest root, say λ_2, of A and its associated vector, say Y, simply because these are the same as the greatest latent root of A_1 and its vector. We may then go on to form a second reduced matrix

$$A_2 = A_1 - \lambda_2 YY' = A - \lambda_1 XX' - \lambda_2 YY',$$

and by the same method find the third greatest root of A; and so on. This process has been referred to as "deflation."

A power of the matrix A obtained to accelerate the convergence may be used to obtain in a very simple manner the same powers of A_1, A_2, and so forth. Let us introduce a matrix

$$S_1 = XX'.$$

Since X has been normalized, $S_1^2 = XX'XX' = XX' = S_1$; and it follows that any positive integral power of S_1 equals S_1 itself. Also, since $AX = \lambda_1 X$ and because the symmetry of A means that $A = A'$, we have

$$AS_1 = AXX' = \lambda_1 XX' = \lambda_1 S_1; \qquad S_1 A = XX'A = \lambda_1 S_1.$$

From these results it follows at once that for every positive integer t

$$A^t S_1 = S_1 A^t = \lambda_1^t S_1 .$$

Since $A_1 = A - \lambda_1 S_1$, it is easy to deduce from these relations that

$$A_1^t = A^t - \lambda_1^t S_1 .$$

Similarly, if $S_2 = YY'$, we find

$$A_2^t = A_1^t - \lambda_2^t S_2 ,$$

and so forth.

A numerical illustration of the application of this method to a set of mental tests was given in the first number of *Psychometrika* [4].

Close and readily calculated limits of error for λ_1 are available. In the last section of [5] it is shown that, for every positive integer t,

$$\lambda_1 \leqq (\operatorname{tr} A^t)^{1/t} ,$$

where "tr" stands for the trace of a matrix, the sum of its diagonal elements. We also have the choice of two lower bounds:

$$\lambda_1 \geqq (\operatorname{tr} A^t/p)^{1/t} ,$$
$$\lambda_1 \geqq X_t' X_t / X_{t-1}' X_t ,$$

where X_t is the vector reached at the tth stage of the iteration. All these bounds converge to λ_1 as t increases.

Limits of error for the components of the latent vectors, and for the latent roots after the first, are not in quite so satisfactory a state. The former have been investigated in the final section just referred to (see also corrections in [6]), with some unexpected help arising from an analogy between

$$\nu_{kt} = X_t' X_{t-k} / X_t' X_t$$

and the kth moment of a probability distribution in which the variate takes only p discrete values equal to the reciprocals of the latent roots. The probability to be associated with the least value of this variate was shown to be a function of the norm of the difference between the true and the obtained vectors X. The Tchebychef inequality then led to an upper bound for this norm. This upper bound is a function of the observable first and second moments, but also of the second greatest root λ_2. This is unfortunate, since we do not yet have entirely satisfactory limits for any of the roots but the greatest. However, a thorough investigation along these lines seems a promising approach to a fully adequate set of limits of error.

One fairly obvious adaptation of the iterative method should perhaps be mentioned. This method determines the roots in the order of their absolute values. If some but not all the roots are wanted, and these not the roots of greatest absolute value, it may still be possible to obtain them by the iterative process without computing the others if a suitable transformation of the matrix can be found. If $f(A)$ is a rational function of a matrix A, its roots are $f(\lambda_i)$, where λ_i are the roots of A, and the latent vectors are the same as those of A. For example, if it is desired to find the two roots of A of least absolute value and their associated vectors, A^{-1} may be calculated and the iterative process applied with it; the two greatest roots of A^{-1} are the reciprocals of the values desired. For A positive definite we might take $f(A) = A - \operatorname{tr} A$.

Four other methods, based on quite different principles, for determination of latent roots and vectors have come to my attention since the publication of the 1943 paper. Limits of error remain to be studied in each case. One of these is the "escalator" method of Morris and Head [**9, 10**], which deals with the more general symmetric determinantal equation $|A - \lambda B| = 0$ and the associated linear equations. The roots and vectors of a matrix are found with the help of those of a submatrix of order less by unity. These are found with the help of roots and vectors of a still smaller submatrix, and so on. The calculation begins by obtaining these quantities for a submatrix of order 2, then for one of order 3 containing the first, and so forth.

An unpublished note by A. R. Collar refers to a method ascribed to Jahn which appears promising. Still different is the earlier method of M. M. Flood [**3**], which uses the Cayley-Hamilton theorem.

Finally, an ingenious method likely to have great value in certain cases was transmitted to me on April 4, 1946, by Rollin F. Bennett. This also uses the Cayley-Hamilton theorem. It requires qualification of the first full paragraph on page 27 of [**5**]. Mr. Bennett's method is as follows:

Let $f_r(\lambda)$ denote the polynomial function $f(\lambda)/(\lambda_r - \lambda)$, where $f(\lambda)$ is the characteristic function of the matrix A, and λ_r is the rth root. Note that $f(A) \equiv (\lambda_1 - A)(\lambda_2 - A) \cdots (\lambda_p - A)$. Since $f(A) \equiv 0$, $(\lambda_r - A)f_r(A) \equiv 0$. This implies that every column of the matrix $f_r(A)$ is proportional to the latent unit vector X_r of A corresponding to λ_r because this vector is defined by

$$(\lambda_r - A)X_r \equiv 0 .$$

Hence, to compute X_r, the following steps suffice:

1. Obtain $A^2, A^3, \cdots, A^{p-1}$ by matrix multiplication.
2. Obtain the coefficients $e_1, \cdots, e_p$ of $f(\lambda)$, from the traces of $A^2, \cdots, A^{p-1}$, using Newton's identities. (Cf. [**5**], p. 24.)
3. Determine the roots $\lambda_1, \cdots, \lambda_p$ of $f(\lambda) = 0$.
4. By division, obtain the coefficients of the reduced polynomial $f_r(\lambda) = f(\lambda)/(\lambda_r - \lambda)$.
5. Compute any column of the matrix $f_r(A)$ by use of the coefficients found in step 4 above and the appropriate columns of the matrices $A, A^2, \cdots, A^{p-1}$ found in step 1.

Convergence of the iterative inversion of a positive definite matrix A can always be assured by taking as the initial approximation C_0 the scalar matrix $(\text{tr } A)^{-1}$, since the latent roots of $D = 1 - AC_0$ then lie between 0 and 1. But the convergence is slow unless a better approximation to A^{-1} is available. The introduction of very high-speed electronic computers capable of multiplication of large matrices in a few minutes reduces the importance of the slowness of convergence. Methods for inversion with such computers are given in a pamphlet by V. Bargmann, D. Montgomery, and J. von Neumann, "Solution of Linear Systems of High Order," a 1946 report under a contract of the Navy Department, Bureau of Ordnance, with the Institute for Advanced Study. An extended discussion of related problems is given by von Neumann and H. H. Goldstine in "Numerical Inverting of Matrices of High Order," *Bulletin of the American Mathematical Society*, vol. 53 (1947), pp. 1021–1099. Other recent notes on matrix calculation and solution of equations are listed in *Mathematical Reviews*. [This paragraph was added in proof.]

REFERENCES

[See Nos. 5 and 13 below for additional references]

1. BERRY, CLIFFORD E. "A criterion of convergence for the classical iterative method of solving simultaneous linear equations," *Annals of Math. Stat.*, vol. 16 (1945), pp. 398–400.
2. DUNCAN, W. J. "Some devices for the solution of large sets of simultaneous linear equations (with an appendix on the reciprocation of partitioned matrices)," *Philos. Mag.*, Ser. 7, vol. 35 (1944), pp. 660–670.
3. FLOOD, MERRILL M. "A computational procedure for the method of principal components," *Psychometrika*, vol. 5 (1940), pp. 169–172.
4. HOTELLING, HAROLD. "Simplified calculation of principal components," *Psychometrika*, vol. 1 (1936), pp. 27–35.
5. ———. "Some new methods in matrix calculation," *Annals of Math. Stat.*, vol. 14 (1943), pp. 1–34.
6. ———. "Further points on matrix calculation and simultaneous equations," *ibid.*, pp. 440–441.
7. HULL, CLARK L. "An automatic correlation calculating machine," *Jour. Amer. Stat. Assoc.*, vol. 20 (1925), pp. 522–531.
8. KOOPMANS, T., and R. B. LEIPNIK. "Iterative computation methods in estimating simultaneous relations," *Econometrica*, vol. 14 (1946), pp. 170–172.
9. MORRIS, J., and J. W. HEAD. "Lagrangian frequency equations. An 'escalator' method for numerical solution" (with an appendix by G. Temple), *Aircraft Engineering*, vol. 14 (1942), pp. 312–314 and 316.
10. ———. "The 'escalator' process for the solution of Lagrangian frequency equations," *Philos. Mag.*, Ser. 7, vol. 35 (1944), pp. 735–759.
11. ULLMAN, JOSEPH. "The probability of convergence of an iterative process of inverting a matrix," *Annals of Math. Stat.*, vol. 15 (1944), pp. 205–213.
12. WAUGH, F. V. "A note concerning Hotelling's method of inverting a partitioned matrix," *Annals of Math. Stat.*, vol. 16 (1945), pp. 216–217.
13. WAUGH, F. V., and P. S. DWYER. "Compact computation of the inverse of a matrix," *Annals of Math. Stat.*, vol. 16 (1945), pp. 259–271.