[3] This argument is to be modified as follows: After the second sentence interpolate "Suppose either X denotes A and Y denotes C or X denotes B and Y denotes F. If X belongs to T let r_X denote XY plus the continuum of the collection G that contains X. If the interval t of the collection T has one end-point at X, let r_T denote $t + x_t + y_t$ where x_t is the continuum of the collection T that contains the other end-point of t and y_t is either XY or XY plus the continuum of G_1 that intersects XY according as XY and G^*_1 are or are not mutually exclusive." In the third sentence, after "T", interpolate "distinct from A and from B." Instead of the sentence beginning in the 18th line of page 381, write "In the space Σ_1, for each continuum q of the collection Q, let g_q denote the boundary of q, plus the set of all points of q, if there are any, which either (a) belong to H or (b) are points of M which in Σ belong to no continuum of the collection H. Let G denote the collection of all such g_q's."

[4] See theorem 16 of chapter V.

[5] Let M denote a continuum formed by a sequence of right pyramids all with the same square base α, their vertices converging to that of the outermost one. Let t_1 and t_2 denote two adjacent sides of α and let H_i denote the set of all intersections of M with planes perpendicular to t_i. Both H_1 and H_2 are arcs with respect to their elements but M is not a continuous curve.

EFFICIENT COMPUTATION OF THE LATENT VECTORS OF A MATRIX

BY PAUL A. SAMUELSON

DEPARTMENT OF ECONOMICS AND SOCIAL SCIENCE, MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Communicated July 8, 1943

In statistics, in quantum mechanics and in the study of dynamical oscillations it is often necessary to compute the latent roots and vectors of a matrix. A variety of methods are available for this purpose, involving direct algebraic computation, iteration and the application of perturbation-variation methods. It is the author's tentative conclusion from experience that the last two methods are excellent if only a few latent roots and vectors are desired, say those corresponding to the lowest or highest few roots, or if there exists some *a priori* familiarity with the data which permits very good initial guesses to be made. But in the general case of high order matrices both the iteration methods and the perturbation-variation methods become tedious, and so recourse must be had to direct algebraic computations.

A variety of methods are available under this heading, and the computer will choose between them not on the basis of their adequacy on constructed text-book examples, but in terms of a careful count of the number of calculations involved in each. Every method must involve the solution of a polynomial of the nth degree; but in addition, all of the methods known to the present writer seem to involve multiplications which increase with the fourth power of n, where the matrix involved is of order n by n. It is the purpose of this note to present, perhaps for the first time, a method which gives latent vectors as well as latent roots after multiplications which increase with the third power of n.

1. Description of Procedure. Let a be the n by n matrix in question, and let h be an arbitrary column vector. Form the matrix products $(h_0, h_1, h_2, \ldots, h_n)$ by means of the operations $[Ih, ah, a(ah), \ldots, a(a^{n-1}h)]$. Then in consequence of the Cayley-Hamilton theorem that a matrix satisfies its own characteristic equation, we derive the coefficients of the characteristic equation $(1, p_1, p_2, \ldots, p_n)$ by solution of the following n linear equations

$$[h_0, h_1, \ldots, h_{n-1}]\begin{bmatrix} p_{n-1} \\ p_{n-2} \\ \cdot \\ \cdot \\ \cdot \\ p_1 \end{bmatrix} = -h_n.$$

Then let the characteristic equation, $f_n(X) = \sum_0^n p_j X^{n-j} = 0$, be solved by any method for the latent roots $(X_1, X_2, \ldots, X_n)$, assumed for simplicity to be distinct. We now form new polynomials by the relations

$$\begin{aligned} f_{n-1}(X) &= X^{n-1} + p_1 X^{n-2} + \ldots + p_{n-1} \\ f_{n-2}(X) &= X^{n-2} + p_1 X^{n-3} + \ldots + p_{n-2} \\ &\cdot \\ &\cdot \\ &\cdot \\ f_1(X) &= X + p_1 \\ f_0(X) &= 1 \end{aligned}$$

where each is formed from the previous by dropping off the last term and lowering the degree of the remaining terms. Expressions of the form $f_i(X_j)$ are easily computed as partial remainders in the familiar process of synthetic division.

Then the n latent vectors of a, $(V_1, V_2, \ldots, V_n)$, can be shown to be given by the product of the following two square matrices

$$V = [h_0, h_1, h_2, \ldots, h_{n-1}]\,[f_{n-i}(X_j)]. \qquad (1)$$

Should the original column vector, h, have been a linear combination of less than n latent vectors, the above process will fail; however, the probability of this occurring is very small, and such occurrences can easily be detected and allowed for. There is no reason why complex latent roots and vectors cannot be handled in the above process. In the important special case where a is symmetrical, only real quantities will be involved, and a check upon the numerical computations is provided by the conjugate property $V_j'V_i = 0$, for $i \neq j$. When repeated roots are encountered, the modifications are relatively minor.

The labor involved in numerical processes of the above type is best reckoned in terms of the required number of multiplications. The method described here involves multiplications of the order $8n^3/3$, or about the equivalent of three square matrix multiplications. In addition, one nth degree polynomial must be solved. It will be noted that each latent vector can be determined independently of all the rest, once its corresponding latent root has been determined. Approximate values of a latent vector can be computed from the insertion of approximate roots in the process of synthetic division indicated above.

2. *Proof*. Consider the system of differential equations written in matrix form

$$DY(t) = aY(t).$$

If the latent roots are all distinct, it is known that the solution of these equations for initial values $Y(0) = h$ is unique and given by

$$\begin{aligned} Y_1(t) &= c_{11} \exp X_1t + c_{12} \exp X_2t + \ldots + c_{1n} \exp X_nt \\ Y_2(t) &= c_{21} \exp X_1t + c_{22} \exp X_2t + \ldots + c_{2n} \exp X_nt \\ &\cdot \\ &\cdot \\ &\cdot \\ Y_n(t) &= c_{n1} \exp X_1t + c_{n2} \exp X_2t + \ldots + c_{nn} \exp X_nt. \end{aligned}$$

The n^2 c coefficients are not all arbitrary, only n of them being dependent upon the initial conditions. Each column of the c's is equivalent to the appropriate latent vector, the initial conditions simply determining the factors of proportionality of the latent vectors.

Our task then is to compute the solution of such a set of differential equations; from this solution we can easily identify the appropriate latent vectors. Thus, we reverse the usual procedure in which the latent vectors are first algebraically computed as an aid in giving the solution of the differential equation system. The novelty of the present method consists in the recognition of this fact plus the specification of a speedy method of arriving at a particular solution of the differential equation system. It is fashionable to handle this last problem by means of the Heaviside-Cauchy operational calculus; a careful consideration of these techniques from a computational point of view will show their efficiency to be greatly overrated, involving in this case multiplications of the order n^4.

It is a commonplace that an nth order differential equation in one variable can be transformed into n first order equations. It is no less true that a system of the latter form can be converted into single equations of the nth order in each variable; for constant coefficient systems such as the one under consideration, the coefficients of the differential equation are in each case simply the coefficients of the characteristic equation; i.e.,

$$f_n(D)Y_j(t) = 0. \qquad (j = 1, \ldots, n).$$

If we can identify the appropriate initial conditions for each of the last equations, and then give the solution of each, we should end up with the required c coefficients indicated above. As for the appropriate initial conditions, if $Y(0) = h$, then by repeated use of the original differential equations, it becomes evident that

$$D^i Y(0) = a^i h = h_i.$$

It is a classical fact that the solution of an nth order differential equation for given initial conditions is given by the solution of a set of linear equations whose matrix is of the Vandermonde-Cauchy form

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ X_1^{n-1} & X_2^{n-1} & \cdots & X_n^{n-1} \end{bmatrix}$$

the transposed inverse of which is given by

$$\begin{bmatrix} \dfrac{P_{n-1}(X_1)}{P_n'(X_1)} & \dfrac{P_{n-1}(X_2)}{P_n'(X_2)} & \cdots & \dfrac{P_{n-1}(X_n)}{P_n'(X_n)} \\ \dfrac{P_{n-2}(X_1)}{P_n'(X_1)} & \dfrac{P_{n-2}(X_2)}{P_n'(X_2)} & \cdots & \dfrac{P_{n-2}(X_n)}{P_n'(X_n)} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \dfrac{1}{P_n'(X_1)} & \dfrac{1}{P_n'(X_2)} & \cdots & \dfrac{1}{P_n'(X_n)} \end{bmatrix}$$

Consequently the n by n c matrix, which is also the matrix made up of columns of latent vectors is given by equation (1) above, where the factors of proportionality $1/P_n'(X_i)$ have been omitted.

When repeated roots are encountered, a generalized Vandermonde determinant is involved whose inversion is easily effected if one simply pursues the close analogy between Vandermonde determinants and expansions in partial fractions.

If the same differential equations are to be solved for many different initial conditions, the above process may be repeated anew. Or a slight economy of effort may be achieved if the inverse of V is worked out once and for all so that the weightings of the different exponential terms can be easily determined by $V^{-1}h$. If the a matrix is symmetrical, the latent vector matrix will be orthogonal so that simple transposition will provide the inverse matrix, except for factors of proportionality.

A SIMPLE METHOD OF INTERPOLATION

By Paul A. Samuelson

Department of Economics and Social Science, Massachusetts Institute of Technology

Communicated July 8, 1943

I.—In many branches of statistics it is necessary to determine the coefficients of an nth degree polynomial, $f(x)$, from $n + 1$ observations, $(x_0, f_0;\ x_1, f_1;\ \ldots;\ x_n, f_n)$, and to determine readings from this polynomial. For this latter purpose recourse may be had to divided differences, Lagrange's interpolation formula,[1] Aitken's method of interpolation,[2] etc. However, where a number of readings are to be taken, or where the coefficients are of interest for their own sake, it is necessary to solve a system of linear equations

$$\begin{array}{l} a_0 + x_0 a_1 + x_0^2 a_2 + \ldots + x_0^n a_n = f_0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_n + x_n a_1 + x_n^2 a_2 + \ldots + x_n^n a_n = f_n \end{array} \quad \text{or } Va = f$$

whose matrix is of the familiar Vandermonde form (x_i^j).

Now in the solution of nth order differential equations with constant coefficients and one-point boundary conditions, such as occur in electrical engineering and other fields of applied mathematics, the solutions can be written in the form of linear combinations of particular solutions, the coefficients being determined by the solution of a transposed Vandermonde set of linear equations.

By means of the Heaviside-Cauchy operational calculus (Laplace transform, etc.), the applied mathematician is able to avoid explicit inversion of such a system of equations. This suggests the possibility of lessening the calculations involved in interpolation by methods analogous to those used in solving differential equations; and upon examination it turns out that the resulting method seems admirably suited to numerical computation, with or without a modern calculating machine.