A METHOD FOR OBTAINING THE CHARACTERISTIC EQUATION OF A MATRIX AND COMPUTING THE ASSOCIATED MODAL COLUMNS*

By HENRY E. FETTIS (*Air Materiel Command*)

The method described below is a process by which the characteristic equation of a matrix may be established without resorting to direct expansion of the coefficients by minors. Further, once the roots of the characteristic equation have been found, the corresponding eigenvectors can be found directly without the additional labor of solving a set of simultaneous equations. Since the only operations involved in the process are the standard ones of matrix multiplication and addition, the work can be set down and performed in routine fashion. The process may also be adapted to punch-card methods for matrices of large order.

1. Preliminary considerations. Consider the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix} \tag{1}$$

The characteristic equation of A is defined as

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{2}$$

When written out in full, (2) is a polynomial equation of degree n in λ:

$$\lambda^n - b_1\lambda^{n-1} + b_2\lambda^{n-2} + \cdots + b_n = 0. \tag{3}$$

The n roots of (3) are called the characteristic numbers or eigenvalues of the matrix A; these numbers are the values of λ for which the homogeneous system of equations

$$\begin{aligned} (a_{11} - \lambda)x_1 + (a_{12})x_2 + \cdots (a_{1n})x_n &= 0, \\ (a_{21})x_1 + (a_{22} - \lambda)x_2 + \cdots (a_{2n})x_n &= 0, \\ \cdots\cdots\cdots\cdots&, \\ (a_{n1})x_1 + (a_{n2})x_2 + \cdots (a_{nn} - \lambda)x_n &= 0, \end{aligned} \tag{4}$$

(or, in matrix form, $Ax = \lambda x$) possesses a non-trivial solution.

*Received Nov. 8, 1949.

The expansion of (2) in the form (3) requires the evaluation of sums of determinants of successively higher order, commencing with b_1 as the sum of the n elements on the principal diagonal. b_1 is called the trace of the matrix A, written

$$b_1 = \text{Tr } A. \tag{5}$$

Similarly, b_2 is the sum of all determinants of order 2 whose diagonal[1] elements coincide with the diagonal elements of A. In all there will be $n(n - 1)/2$ such determinants. In a like manner, b_3 is made up of the sum of the $n(n - 1)(n - 2)/6$ determinants of order 3 which can be formed in this way; continuing, it is found that $b_n = |A|$. It is evident that as the order of the matrix increases, the labor involved in expanding the characteristic equation also increases, but much more rapidly, so that a point is soon reached for which direct expansion is impractical.

In addition to the work required in deriving the characteristic equation, the complete solution of the problem will in general require the resubstitution of the roots of the equation into the original system of equations to obtain the corresponding relation between the x_i. Since for λ equal to any root of (3), the system (4) becomes consistent, it suffices to solve for $(n - 1)$ of the x_i in terms of any arbitrary one. The relation between the x_i which is now determined apart from an arbitrary numerical factor (whose actual value is usually of no concern) is called the modal column or eigenvector corresponding to the characteristic number. Obviously for a matrix of order 4 or higher such a solution would become extremely laborious in view of the fact that for each of the n characteristic numbers, a system of $(n - 1)$ simultaneous equations must be solved.

In the following section a method is set forth by which the characteristic equation of a matrix may be established in a routine manner employing the standard operations of matrix multiplication and addition. The labor does not increase with the order of the matrix to the extent that direct expansion does. In addition the results so obtained eliminate the necessity of solving a set of simultaneous equations for each root of the characteristic equation.

2. Summary of the method. Let the given matrix be A, and define the successive matrices A_k and numbers b_k as follows:

$$\begin{aligned} A_0 &= A, & b_1 &= (\text{Tr } A_0); \\ A_1 &= b_1 A_0 - A_0^2, & b_2 &= (\text{Tr } A_1)/2; \\ A_2 &= b_2 A_0 - A_0 A_1, & b_3 &= (\text{Tr } A_2)/3; \end{aligned} \tag{6}$$

in general

$$A_k = b_k A_0 - A_0 A_{k-1}, \qquad b_{k+1} = (\text{Tr } A_k)/(k + 1).$$

Continuing the above, it will be found that $A_n = 0$. This result serves as a check on the correctness of the operations. When this point has been reached, the characteristic equation is

$$\lambda^n - b_1\lambda^{n-1} + b_2\lambda^{n-2} + \cdots + (-)^n b_n = 0. \tag{7}$$

[1]The term "diagonal elements" is used to denote the elements of the principal diagonal.

Further, if λ_i is any root of (7), the corresponding eigenvector $x^{(i)}$ is proportional to any column of the matrix

$$A_0\lambda_i^n - A_1\lambda_i^{n-1} + A_2\lambda_i^{n-2} + \cdots (-)^{n-1}A_{n-1}\lambda_i \ . \tag{8}$$

In actual practice it is not necessary to compute all the elements of the various matrices. In fact, if $n \geq 4$, only $(n - 3)$ of the matrices A_k need be computed in their entirety.

The method and its modifications are illustrated in the next section.

3. Numerical example[2]. For the matrix

$$A = \begin{bmatrix} -2 & -2 & 0 & 3 & -1 \\ -2 & 0 & -3 & 5 & 0 \\ 0 & -3 & -5 & 1 & 1 \\ 3 & 5 & 1 & -3 & -1 \\ -1 & 0 & 1 & -1 & -1 \end{bmatrix}$$

we have

$$A_0 = \begin{bmatrix} -2 & -2 & 0 & 3 & -1 \\ -2 & 0 & -3 & 5 & 0 \\ 0 & -3 & -5 & 1 & 1 \\ 3 & 5 & 1 & -3 & -1 \\ -1 & 0 & 1 & -1 & -1 \end{bmatrix} \quad \begin{vmatrix} -2 \\ 0 \\ -6 \\ 5 \\ -2 \end{vmatrix}$$

$$A_0^2 = \begin{bmatrix} 18 & 19 & 8 & -24 & 0 \\ 19 & 38 & 20 & -24 & -6 \\ 8 & 20 & 36 & -24 & -7 \\ -24 & -24 & -24 & 45 & 2 \\ 0 & -6 & -7 & 2 & 4 \end{bmatrix} \quad \begin{vmatrix} 21 \\ 47 \\ 33 \\ -25 \\ -7 \end{vmatrix}$$

Thus, $b_1 = -11$.

(The columns on the right consist of the sums of the elements of the respective rows. These serve as a check on the matrix multiplication, since the product of the sum column

[2]This example is taken from W. Kincaid's paper *Numerical methods for finding the characteristic roots and vectors of matrices*, Q. Appl. Math, 5, 320-345 (1947).

of any matrix into any other matrix should equal the sum-column of the product matrix.)

Continuing we compute $A_1 = b_1 A_0 - A_0^2$ and $A_0 A_1$:

$$A_1 = \left[\begin{array}{ccccc|c} 4 & 3 & -8 & -9 & 11 & 1 \\ 3 & -38 & 13 & -31 & 6 & -47 \\ -8 & 13 & 19 & 13 & -4 & 33 \\ -9 & -31 & 13 & -12 & 9 & -30 \\ 11 & 6 & -4 & 9 & 7 & 29 \end{array}\right]; \quad A_0 A_1 = \left[\begin{array}{ccccc|c} -52 & -29 & 33 & 35 & -14 & -27 \\ -29 & -200 & 24 & -81 & 35 & -251 \\ 33 & 24 & -125 & 25 & 18 & -25 \\ 35 & -81 & 25 & -142 & 25 & -138 \\ -14 & 35 & 18 & 25 & -31 & 33 \end{array}\right]$$

Thus $b_2 = -20/2 = -10$. Next we find A_2 from $A_2 = b_2 A_0 - A_0 A_1$, and the diagonal elements only of the product $A_0 A_2$.

$$A_2 = \left[\begin{array}{ccccc|c} 72 & 49 & -33 & -65 & 24 & 47 \\ 49 & 200 & 6 & 31 & -35 & 251 \\ -33 & 6 & 175 & -35 & -28 & 85 \\ -65 & 31 & -35 & 172 & -15 & 88 \\ 24 & -35 & 26 & -15 & 41 & -13 \end{array}\right]; \quad A_0 \cdot A_2 = \begin{bmatrix} -461 & & & & \\ & 39 & & & \\ & & -956 & & \\ & & & -576 & \\ & & & & -78 \end{bmatrix}.$$

Thus, $b_3 = 660/3 = 220$. The diagonal elements of A_3 and the coefficient b_4 may now be found from $A_3 = b_3 A_0 - A_0 A_2$ and $b_4 = \text{Tr } A_3$. The expression for A_4 is rewritten as

$$A_4 = b_4 A_0 - (b_3 A_0^2 - A_0^2 A_2).$$

Again only the diagonal elements are needed.

$$A_3 = \begin{bmatrix} 21 & & & & \\ & -39 & & & \\ & & -144 & & \\ & & & -84 & \\ & & & & -142 \end{bmatrix}; \quad A_0^2 A_2 = \begin{bmatrix} 3523 & & & & \\ & 8117 & & & \\ & & 7192 & & \\ & & & 9366 & \\ & & & & 540 \end{bmatrix}.$$

Thus, $b_4 = -388/4 = -97$.

$$A_0 \cdot A_3 = b_3 A_0^2 - A_0^2 \cdot A_2 = \begin{bmatrix} 433 & & & & \\ & 243 & & & \\ & & 728 & & \\ & & & 534 & \\ & & & & 340 \end{bmatrix};$$

$$A_4 = b_4 A_0 - A_0 A_3 = \begin{bmatrix} -243 & & & & \\ & -243 & & & \\ & & -243 & & \\ & & & -243 & \\ & & & & -243 \end{bmatrix}.$$

Hence $b_5 = -243$ and the characteristic equation is

$$\lambda^5 + 11\lambda^4 - 10\lambda^3 + 220\lambda^2 - 97\lambda + 243 = 0.$$

The roots[3] of this equation are

$\lambda_1 = -9.886487, \quad \lambda_2 = -4.75775, \quad \lambda_3 = 4.22365, \quad \lambda_4 = -1.43300, \quad \lambda_5 = .85355.$

Rewriting the first columns of A_0, A_1, A_2 and computing the first columns of A_3 and A_4 by column multiplication gives

$$(A_0) = \begin{bmatrix} -2 \\ -2 \\ 0 \\ 3 \\ -1 \end{bmatrix}; \quad (A_1) = \begin{bmatrix} 4 \\ 3 \\ -8 \\ -9 \\ 11 \end{bmatrix}; \quad (A_2) = \begin{bmatrix} 72 \\ 49 \\ -33 \\ -65 \\ 24 \end{bmatrix}; \quad (A_3) = b_3(A_0) - (A_0)(A_2) = \begin{bmatrix} 21 \\ -70 \\ 23 \\ 61 \\ -156 \end{bmatrix};$$

where (A) denotes the first column of A. Finally

[3]A convenient method for solving equations of this type is described by Shih-Nge Lin, *A method of successive approximations of evaluating the real and complex roots of cubic and higher equations*, J. Math. Phys., **20**, 231-242 (1941).

$$(A_4) = b_4(A_0) - A_0(A_3) = \begin{bmatrix} -243 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

As a further check, it is noted that the elements of (A_4) not previously computed are zero.

The eigenvectors are therefore (after removing a factor of λ_i) proportional to

$$\begin{bmatrix} -2 \\ -2 \\ 0 \\ 3 \\ -1 \end{bmatrix} \lambda_i^4 - \begin{bmatrix} 4 \\ 3 \\ -8 \\ -9 \\ 11 \end{bmatrix} \lambda_i^3 + \begin{bmatrix} 72 \\ 49 \\ -33 \\ -65 \\ 24 \end{bmatrix} \lambda_i^2 - \begin{bmatrix} 21 \\ -70 \\ 23 \\ 61 \\ -156 \end{bmatrix} \lambda_i + \begin{bmatrix} -243 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

A convenient method of evaluating the above polynomials is by synthetic division. Thus for $\lambda_1 = -9.886487$,

$$\begin{array}{rrrrr}
-2 & -\ 4 & +\ 72 & -\ 21 & -\ 243 \\
 & 19.772974 & -\ 155.9393 & +\ 829.8648 & -\ 7996.83 \\ \hline
-2 & +\ 15.772974 & -\ 83.9393 & +\ 808.8648 & \boxed{-\ 8239.83}
\end{array}$$

$$\begin{array}{rrrrr}
-2 & -\ 3 & +\ 49 & +\ 70 & +\ 0 \\
 & 19.772974 & -\ 165.8258 & +\ 1154.997 & -\ 12110.917 \\ \hline
-2 & +\ 16.772974 & -\ 116.8258 & +\ 1224.997 & \boxed{-\ 12110.917}
\end{array}$$

$$\begin{array}{rrrrr}
0 & +\ 8 & -\ 33 & -\ 23 & +\ 0 \\
 & 0 & -\ 79.0919 & +\ 1108.195 & -\ 10728.766 \\ \hline
0 & +\ 8 & -\ 112.0919 & +\ 1085.195 & \boxed{-\ 10728.766}
\end{array}$$

$$\begin{array}{rrrrr}
3 & +\ 9 & -\ 65 & -\ 61 & +\ 0 \\
 & -\ 29.65946 & +\ 204.2495 & -\ 1376.688 & +\ 14213.68 \\ \hline
3 & -\ 20.65946 & +\ 139.2495 & -\ 1437.688 & \boxed{+\ 14213.68}
\end{array}$$

$$\begin{array}{rrrrr}
-1 & -\ 11 & +\ 24 & +\ 156 & +\ 0 \\
 & +\ 9.886487 & +\ 11.00873 & -\ 346.11335 & +\ 1879.553 \\ \hline
-1 & -\ 1.113513 & +\ 35.00873 & -\ 190.11335 & \boxed{+\ 1879.553}
\end{array}$$

Therefore the first eigenvector is proportional to

$$\begin{bmatrix} -8239.83 \\ -12110.92 \\ -10728.77 \\ 14213.68 \\ 1879.55 \end{bmatrix}$$

or to

$$x^{(1)} = \begin{bmatrix} 1.000000 \\ 1.469802 \\ 1.302062 \\ -1.724997 \\ -\ .228106 \end{bmatrix}.$$

The remaining eigenvectors, found in a similar manner are:

$$x^{(2)} = \begin{bmatrix} 1.00000 \\ .08337 \\ -1.0793 \\ -\ .6869 \\ .5302 \end{bmatrix};\ x^{(3)} = \begin{bmatrix} .04830 \\ 1.00000 \\ -\ .2703 \\ .7018 \\ -\ .1935 \end{bmatrix};\ x^{(4)} = \begin{bmatrix} .4224 \\ -\ .00483 \\ .3993 \\ .4099 \\ 1.0000 \end{bmatrix};\ x^{(5)} = \begin{bmatrix} 1.0000 \\ -\ .4364 \\ .1825 \\ .4350 \\ -\ .6757 \end{bmatrix}.$$

In this example, direct expansion of the characteristic equation would involve the computation of 10 determinants of order 2, 10 of order 3, 5 of order 4, and 1 of order 5. In addition the only available check would be an independent evaluation. Further, for each characteristic number, the solution of a system of four equations would be required, necessitating the additional evaluation of 5 determinants of order 4 for each λ_i .