L. M. MILNE-THOMSON, *The Calculus of Finite Differences.* Macmillan, London, 1933. 558 p.

N. E. NÖRLUND, *Vorlesungen über Differenzenrechnung.* (Reprint) Edwards, Ann Arbor, 1945. 551 p.

W. SCHMEIDLER, *Integralgleichungen mit Anwendungen in Physik und Technik.* I. *Lineare Integralgleichungen.* Akademische Verlagsgesellschaft, Leipzig, 1950. 611 p.

o. *Inequalities*

G. H. HARDY, J. E. LITTLEWOOD, & G. PÓLYA, *Inequalities.* Cambridge University Press, 1934. 314 p.

T. MOTZKIN, *Beiträge zur Theorie der linearen Ungleichungen.* Azriel, Jerusalem, 1936. 71 p. [Translated into English by D. R. FULKERSON, Publication T-22, The RAND Corporation, Santa Monica, Calif., 7 March 1952, 86 p.]

p. *Calculus of variations, etc.*

G. A. BLISS, *Lectures on the Calculus of Variations.* University of Chicago Press, 1946. 296 p.

R. COURANT & D. HILBERT, *Methoden der mathematischen Physik.* 2d edition, 2 vols., reprint, Interscience, New York, 1943. 469 + 549 p.

q. *Probability*

H. CRAMÉR, *Mathematical Methods of Statistics.* Princeton University Press, 1946. 575 p.

W. FELLER, *An Introduction to Probability Theory and its Applications.* v. 1, Wiley, New York, 1950. 419 p.

J. V. USPENSKY, *Introduction to Mathematical Probability.* McGraw-Hill, New York, 1937. 411 p.

r. *Topology*

M. H. A. NEWMAN, *Elements of the Topology of Plane Sets of Points.* 2nd edition, Cambridge University Press, 1951. 214 p.

s. *Geometry*

P. J. KELLY & H. BUSEMANN, *Projective Geometry and Projective Metrics.* Academic Press, 1953. 332 p.

I. S. SOKOLNIKOFF, *Tensor Analysis, Theory and Applications.* Wiley, 1951. 335 p.

D. J. STRUIK, *Lectures on Classical Differential Geometry.* Addison-Wesley, Cambridge, 1950. 221 p.

NBSINA

GEORGE E. FORSYTHE

This compilation was sponsored in part by the Office of Naval Research, USN. It represents the opinions of the author only, and not those of the National Bureau of Standards or the Office of Naval Research.

On Finding the Characteristic Equation of a Square Matrix

Various methods are known for finding explicitly the characteristic equation of a square matrix.[1,2,3,4,5,6] Some of these make use of the Cayley-Hamilton theorem which states that every square matrix satisfies its own characteristic equation.[3,4,5] In the present paper we describe among others, the method of K. HESSENBERG,[1,2] who uses the fact that similar matrices have the same characteristic equation. We also indicate a useful combination of Hessenberg's method with an iteration technique based upon the Cayley-Hamilton theorem. The Hessenberg method is as follows:

Let A be the given $n \times n$ matrix whose characteristic equation is sought. With an arbitrary column vector z_1 and certain scalars p_{ij}, to be defined presently, form the columns $z_2, z_3, \cdots, z_{n+1}$ in the manner

$$
\begin{aligned}
z_2 &= Az_1 + p_{11}z_1 \\
z_3 &= Az_2 + p_{12}z_1 + p_{22}z_2 \\
z_4 &= Az_3 + p_{13}z_1 + p_{23}z_2 + p_{33}z_3 \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
z_{n+1} &= Az_n + p_{1n}z_1 + p_{2n}z_2 + \cdots + p_{nn}z_n.
\end{aligned} \tag{1}
$$

Choose the p_{ij} so that the matrix Z has the form

$$
Z = (z_1, z_2, \cdots, z_n) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & z_{22} & 0 & \cdots & 0 \\ 0 & z_{32} & z_{33} & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & z_{n2} & z_{n3} & \cdots & z_{nn} \end{bmatrix} \tag{2}
$$

that is, the p_{ij} are taken such that z_1 has unity for its first component and zero for each of its other components, while the first k components of z_{k+1} vanish for $k = 1, 2, \cdots n - 1$.

Thus, defining the matrix P as

$$
P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1,n-1} & p_{1n} \\ -1 & p_{22} & p_{23} & \cdots & p_{2,n-1} & p_{2n} \\ 0 & -1 & p_{33} & \cdots & p_{3,n-1} & p_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & -1 & p_{nn} \end{bmatrix}, \tag{3}
$$

Hessenberg finds

$$AZ + ZP = 0. \tag{4}$$

For

$$AZ = (Az_1, Az_2, \cdots, Az_n) \tag{5}$$

and if in (5) we replace $Az_1, Az_2, \cdots Az_n$ by their equivalents as found from (1), we have

$$AZ = (z_2 - p_{11}z_1, z_3 - p_{12}z_1 - p_{22}z_2, z_4 - p_{13}z_1 - \cdots - p_{33}z_3, \cdots). \tag{6}$$

On the other hand,

$$
\begin{aligned}
ZP &= z_1(p_{11}, p_{12}, \cdots p_{1n}) + z_2(-1, p_{22}, \cdots, p_{2n}) + \cdots \\
&= (p_{11}z_1, p_{12}z_1, \cdots, p_{1n}z_1) + (-z_2, p_{22}z_2, \cdots p_{2n}z_2) + \cdots
\end{aligned} \tag{7}
$$

and on adding (6) and (7) we obtain (4). Notice by (2) that the columns $z_1, z_2, \cdots, z_n$ are, in general, linearly independent. Thus when Z^{-1} exists, from (4) it follows that

$$-P = Z^{-1}AZ$$

which shows that $-P$ is similar[6] to A, and by a known theorem[6] it follows that $-P$ has the same characteristic equation as A.

The characteristic equation is

$$\det\,(P + \lambda I) = 0, \text{ i.e.,}$$

$$\begin{vmatrix} (\lambda + p_{11}) & +p_{12} & +p_{13} & \cdots & +p_{1n} \\ -1 & (\lambda + p_{22}) & +p_{23} & \cdots & +p_{2n} \\ 0 & -1 & (\lambda + p_{33}) & \cdots & +p_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & (\lambda + p_{nn}) \end{vmatrix} = 0.$$

Expanding the determinant by the elements of its last column, one finds that the equation is

$$p_{1n} + p_{2n}F_1 + p_{3n}F_2 + \cdots + p_{n-1,n}F_{n-2} + p_{nn}F_{n-1} = 0, \tag{8}$$

where the F's may be calculated by recursion:

$$\begin{aligned} F_1 &= \lambda + p_{11} \\ F_2 &= (\lambda + p_{22})F_1 + p_{12} \\ F_3 &= (\lambda + p_{33})F_2 + p_{23}F_1 + p_{13} \\ F_4 &= (\lambda + p_{44})F_3 + p_{34}F_2 + p_{24}F_1 + p_{14} \\ &\;\;\vdots \\ &\text{etc.} \end{aligned} \tag{9}$$

The matrices Z and P can be computed systematically.

Form the array:

$$\begin{array}{|c|c|} \hline A & Z \\ \hline \multicolumn{1}{c|}{} & P \\ \cline{2-2} \end{array}$$

It follows from (4) that the scalar product of any row of (A, Z) by any column of $\begin{bmatrix} Z \\ P \end{bmatrix}$ must vanish. If these scalar products are formed in proper succession, setting such a product equal to zero gives a simple equation containing just one "unknown", a z_{ij} or a p_{ij}, which is readily determined. Thus:

Multiplication of the first column of $\begin{pmatrix} Z \\ P \end{pmatrix}$

by the 1st row of (A, Z) permits determination of p_{11}
by the 2nd row of (A, Z) permits determination of z_{22}
by the 3rd row of (A, Z) permits determination of z_{32}
by the 4th row of (A, Z) permits determination of z_{42}
etc. etc.

Multiplication of the second column of $\begin{pmatrix} Z \\ P \end{pmatrix}$

by the 1st row of (A, Z) permits determination of p_{12}
by the 2nd row of (A, Z) permits determination of p_{22}
by the 3rd row of (A, Z) permits determination of z_{33}
by the 4th row of (A, Z) permits determination of z_{43}
etc. etc.

Similarly, multiplication of each of the other columns of $\begin{pmatrix} Z \\ P \end{pmatrix}$ in proper succession by each of the rows of (A, Z) permits determination of all remaining elements.

After thus determining matrix P, the F's may be computed from (9) and the characteristic equation by (8).

Exceptional Cases. As pointed out by ZURMÜHL,[2] minor difficulties sometimes occur. For instance, one of the z_{kk} may vanish when, according to the given procedure, it is used as a divisor to determine some element of P. Also an entire column of z_{ij}'s may consist of zeros. Treatment of such cases is indicated in the following examples.[2]

Case 1. Vanishing of a z_{kk}.

In Fig. 1 the work proceeds normally until we try to find p_{22}, where we have $2 + z_{22}p_{22} = 0$ which is meaningless since $z_{22} = 0$. However, this difficulty can be avoided in a way consistent with (4).

$$\begin{array}{rrr|rrr} 2 & 3 & -2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 2 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ \hline & & & -2 & 2 & -6 \\ & & & -1 & 1 & -4 \\ & & & 0 & -1 & -1 \end{array}$$

FIG. 1.

Assign z_{33} arbitrarily, say $z_{33} = 0$, and let z_{23} be "unknown." Then: Equating to zero the product of

Row 3 of (A, Z) by column 2 of $\binom{Z}{P}$, determines $p_{22} = 1$,
Row 2 of (A, Z) by column 2 of $\binom{Z}{P}$, determines $z_{23} = 2$,
Row 1 of (A, Z) by column 3 of $\binom{Z}{P}$, determines $p_{13} = -6$,
Row 3 of (A, Z) by column 3 of $\binom{Z}{P}$, determines $p_{23} = -4$,
Row 2 of (A, Z) by column 3 of $\binom{Z}{P}$, determines $p_{33} = -1$.

Since the array of Fig. 1 now satisfies (4), and P has the standard form, we apply (9) and (8) to find the characteristic equation

$$\lambda^3 - 2\lambda^2 - 3\lambda + 2 = 0.$$

Alternatively, we could have assigned p_{22} arbitrarily, leaving z_{23}, z_{33}, p_{13}, p_{23}, p_{33} to be determined so that (4) holds. The resulting P will be similar to the one already found and the characteristic equation will be unchanged.

Case 2. Vanishing of a vector z_k.

In Fig. 2 it happens that $z_3 = 0$, apparently indicating that $p_{13} = p_{23} = p_{33} = 0$. Here we replace z_3 by ϵz_3 and let ϵ tend to zero in the final calculation.

$$\begin{array}{rrr|rrr} -3 & 1 & 3 & 1 & 0 & 0 \\ 10 & 0 & -6 & 0 & 10 & 0 \\ -10 & 2 & 8 & 0 & -10 & \epsilon \\ \hline & & & 3 & 20 & -3\epsilon \\ & & & -1 & -6 & 0.6\epsilon \\ & & & 0 & -1 & -2 \end{array}$$

FIG. 2.

Thus we find $p_{13} = -3\epsilon$, $p_{23} = 0.6\epsilon$, $p_{33} = -2$, and

$$F_1 = \lambda + 3$$
$$F_2 = (\lambda - 6)F_1 + 20 = \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$$
$$F_3 = (\lambda - 2)F_2 + 0 + 0 = (\lambda - 2)^2(\lambda - 1) = 0.$$

Evidently Case 2 occurs when the characteristic equation has a multiple root and A has linear elementary divisors corresponding to the root.

Iteration Techniques. While the use of (9) and (8) is a highly efficient way of getting the characteristic equation (8), it would appear that (9) and (8) are not as automatic as the method of finding P. To obtain a more fully automatic procedure, we propose that the Cayley-Hamilton theorem[6] be applied to the Hessenberg matrix P in the following way.

Let the characteristic equation of P be

$$c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_{n-1}\lambda^{n-1} + \lambda^n = 0 \tag{10}$$

Then from the Cayley-Hamilton equation it follows that

$$c_0I + c_1P + c_2P^2 + \cdots + P^n = 0. \tag{11}$$

Post-multiply (11) by an arbitrary column matrix x_0. There results

$$c_0x_0 + c_1x_1 + c_2x_2 + \cdots + x_n = 0 \tag{12}$$

where

$$P^kx_0 \equiv x_k. \tag{13}$$

If x_0 is taken to be the column matrix $\{1, 0, 0, \cdots, 0\}$, then (12) becomes in general a triangular system of linear equations from which the c's are found in easy succession. The desired equation is finally obtained from (10) on replacing λ by $-\lambda$. If we define the matrices

$$X = (x_0, x_1, \cdots, x_n),$$

$$C = \begin{bmatrix} c_0 \\ c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_{n-1} \end{bmatrix}$$

our complete scheme is represented by the array:

$$\begin{array}{|c|c|c|c|}\cline{1-2} A & Z & \multicolumn{2}{c}{} \\ \hline \multicolumn{1}{c|}{} & P & X & C \\ \cline{2-4}\end{array} \tag{14}$$

Thus, the procedure is fully automatic.

If the given matrix A has a sufficient number of zero elements in its lower left corner, and in particular if A is a continuant, then Hessenberg's Z and P are unnecessary in our method since we obtain a much more rapid solution by direct application of the iteration technique, as indicated by the array:

$$\begin{array}{|c|c|c|}\hline A & X & C \\ \hline\end{array} \tag{15}$$

In the cases where (15) applies, (10) directly represents the characteristic equation of A, so that here we need not replace λ by $-\lambda$.

Elementary Transformations. The elementary transformations or operations[6] which, applied either singly or jointly to the square matrix A, leave invariant the characteristic equation of A are as follows:

1. If the i-th and j-th rows are interchanged, the i-th and j-th columns must be interchanged.
2. If the elements of the i-th row are multiplied by k, the elements of the i-th column must be multiplied by $1/k$.
3. If k times the j-th row is added to the i-th row, then the negative of k times the i-th column must be added to the j-th column.

Using such elementary operations, we can always transform A so that it has an isosceles right-triangular array of $(n-1)(n-2)/2$ ($\equiv$ triangular number of order $n-2$) zeros in its lower left corner, where n is the order of A, so that by using the scheme (15) on the new matrix we readily obtain the characteristic equation of A.

The presence of zero elements in A generally facilitates the transformation. Generally the transformation is not too laborious in any specific case and we recommend that the characteristic equation be found by applying scheme (15) to the transformed matrix rather than by applying scheme (14) to the original matrix. A great deal will of course depend on the nature of the matrix and the kind of equipment available for computation.

Examples. Consider the matrix A_1,

$$A_1 = \begin{bmatrix} 2 & 3 & -2 \\ 0 & 1 & 2 \\ 1 & 2 & -1 \end{bmatrix}$$

which was used in a previous illustration and with which there was some difficulty using the Hessenberg method.

By interchanging the second and third rows and then interchanging the second and third columns of A_1, we obtain

$$T_1 = \begin{bmatrix} 2 & -2 & 3 \\ 1 & -1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

On applying scheme (15) to T_1, we have

2	−2	3	1	2	2	8	2
1	−1	2	0	1	1	5	−3
0	2	1	0	0	2	4	−2

From the last column, we read that the characteristic equation of A_1 is

$$\lambda^3 - 2\lambda^2 - 3\lambda + 2 = 0.$$

An example of the complete procedure (14) for a matrix[3] of fourth order is the following:

1	−2	3	−2	1	0	0	0						
1	5	−1	−1	0	1	0	0						
2	3	2	−2	0	2	1	0						
2	−2	6	−3	0	2	2	−2						
				−1	0	1	−4	1	−1	1	0	−2	−8
				−1	−1	3	−2	0	−1	2	0	−16	−4
				0	−1	−4	0	0	0	1	−6	24	6
				0	0	−1	1	0	0	0	−1	5	5

From the column C we infer

$$\det (P - I\lambda) = \lambda^4 + 5\lambda^3 + 6\lambda^2 - 4\lambda - 8 = 0.$$

Hence the characteristic equation of A is

$$\det (A - \lambda I) = \det (P + I\lambda) = \lambda^4 - 5\lambda^3 + 6\lambda^2 + 4\lambda - 8 = 0.$$

Number of Operations Required. An "operation," for the present purpose, is either a multiplication or a division of a pair of numbers.

1. (a) If we ignore multiplication by λ, such as $\lambda \cdot p_{22}$ etc., in the formation of the F's, then to obtain the characteristic equation Hessenberg requires $M_1 = n^3 - (3/2)n^2 + (1/2)n$ multiplications and $D_1 = n(n-1)/2$ divisions.
   (b) If we include trivial multiplications such as $\lambda \cdot p_{22}$, then Hessenberg requires $M_2 = n^3 - (1/2)n^2 + (3/2)n$ multiplications and D_1 divisions.
2. If we use Hessenberg's method to obtain P and then apply the Cayley-Hamilton theorem to P, then to obtain the characteristic equation in this way we require $M_3 = n^3 - n^2$ multiplications and D_1 divisions.
3. If the matrix A has no zeros that may be taken advantage of, and elementary operations are used to obtain a similar matrix with $(n-1)(n-2)/2$ zeros in the lower left corner, and if the iteration scheme (15) is used starting with the column vector $(1, 0, 0, \cdots, 0)$, then $(7/6)n^3 - 2n^2 + 17n/6 - 3$ multiplications and n divisions are required to find the characteristic equations. This indicates that, at least for large n, methods 1, 2, and 3 require nearly the same

number of operations, while method 2 has the advantage of being most automatic. However, for special matrices or matrices having many zeros method 3 may be best.

4. The methods of FRAME[3] or HOTELLING-BINGHAM-GIRSHICK[3] require a number of operations of the order of n^4 since they depend upon $n - q$ multiplications of n by n matrices. Multiplication of two n by n matrices requires n^3 operations, and q is some constant independent of n, usually 1, 2, or 3.

Remarks. A use for the characteristic equation which does not seem to have been explicitly mentioned in the literature is the following. In many physical problems it is necessary to find the eigenvalues and eigenvectors of a square non-singular matrix. For this purpose iterative methods are generally applied. Quite often it is sufficient to find not all of the eigenvalues but only a few of the lowest ones. Since the method of iteration leads to the dominant eigenvalue, it has been necessary to find the inverse matrix to use the fact that the largest eigenvalue of the inverse matrix is the reciprocal of the smallest eigenvalue of the original matrix.

However, instead of finding the inverse of the original matrix, one may find the characteristic equation of the original matrix explicitly and at once write down the equation which has for its roots the reciprocals of the original roots. If (10) is the characteristic equation then

$$\mu^n + \frac{c_1}{c_0}\mu^{n-1} + \cdots + \frac{c_{n-1}}{c_0}\mu + \frac{1}{c_0} = 0 \tag{16}$$

has the reciprocals of λ as its roots, which may be found by any convenient method.

The roots of (16) may also be found by matrix iteration methods applied to the *comparison matrix* or *companion matrix:*

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ -\frac{1}{c_0} & -\frac{c_{n-1}}{c_0} & \cdots & & -\frac{c_1}{c_0} \end{bmatrix}$$

which has (16) as its characteristic equation. Inversion of the original matrix is thus avoided, and incidentally the iteration procedure is easily carried out because of the large number of zeros in the matrix.

It is interesting to note that DANILEVSKI[4] has devised a method for finding the characteristic equation by reducing the matrix A to the companion matrix form by elementary transformations. However, this method does not compare favorably with the first three discussed above.

In conclusion we should like to point out that some of the difficulties encountered by Hessenberg are avoided by the method of obtaining zeros in the lower left hand corner and using iteration. One such example has already been shown. Consider now the other example, which had to be treated as a special case. If row 2 is added to row 3 and then column 3 is subtracted from column 2, the matrix

$$A_2 = \begin{bmatrix} -3 & 1 & 3 \\ 10 & 0 & -6 \\ -10 & 2 & 8 \end{bmatrix}$$

becomes the similar matrix

$$T_2 = \begin{bmatrix} -3 & -2 & 3 \\ 10 & 6 & -6 \\ 0 & 0 & 2 \end{bmatrix}$$

Here obviously $(\lambda - 2)$ is a factor and all that remains to do is find the other factor from the matrix

$$\begin{bmatrix} -3 & -2 \\ 10 & 6 \end{bmatrix}$$

This is easily found to be $(\lambda - 2)(\lambda - 1)$.

EDWARD SAIBEL
W. J. BERGER

Carnegie Institute of Technology
Pittsburgh, Pa.

This work was supported in part by a contract between Carnegie Institute of Technology and the Department of the Army Ordnance Corps.

[1] FIAT *Review of German Science*, Applied Mathematics. Part I, 1948, p. 31–33.
[2] R. ZURMÜHL, *Matrizen*. Berlin, 1950, p. 316–322.
[3] P. S. DWYER, *Linear Computations*. New York, 1951.
[4] I. M. GEL'FAND, *Lekt͡sii po Lineĭnoĭ Algebre* [*Lectures on Linear Algebra*]. Moscow 1951, p. 233–239.
[5] H. E. FETTIS, "A method for obtaining the characteristic equation of a matrix and computing the associated modal columns," *Quart. Appl. Math.*, v. 8, 1950, p. 206–212.
[6] G. BIRKHOFF & S. MACLANE, *A Survey of Modern Algebra*. New York, 1941.

RECENT MATHEMATICAL TABLES

1122[A,B,C].—P. P. ANDREEV, *Matematicheskie Tablit͡sy* [Mathematical Tables]. Moscow, 1952, 471 p. 12.5 × 19.7 cm. Price 7.75 rubles.

The main table of this work is a table of n^s for $n = 1(1)10000$, $s = -1/2$, 3, 2, 1/2, 1/3. Values are given to 6S only. The values of $(10n)^{1/2}$ are also given for $n > 1000$ while the natural logarithm of n is given for $n \leq 1000$. This part occupies 333 pages, only 30 values of n being devoted to each page. This table is certainly no substitute for BARLOW.

The second part of the volume is devoted to 19 small tables of minor importance including a 6S table of $1/n$ for $n = 1(1)10000$, log n for $n = 1(1)1000$ to 9D, and the binomial coefficients of the first 50 integral powers.

D. H. L.

1123[A].—M. LOTKIN & M. E. YOUNG, *Table of Binomial Coefficients*. Ballistic Research Laboratories *Memo. Report* No. 652. Aberdeen Proving Ground, 1953, 37 p., 21.6 × 27.9 cm. mimeographed from typescript.