

On the strength of relations (9), the coefficients p_1, p_2, \dots, p_n are non-homogeneous linear forms in q_1, \dots, q_{n-1} and consequently may be calculated simultaneously by the method of elimination (see § 12). Of the two possible modifications of the method of elimination, one should take that in which the components of the vectors $R', M'R', \dots, M'^{n-1}R'$, are arrayed in the rows of the scheme. Then these rows, conceived as matrices, are R, RM, \dots, RM^{n-1} . The coefficients of the relations (9) will thereupon turn out to be arranged so as to accord exactly with the Samuelson scheme. Given the grounds for the method that have been presented above, the region of its application is easily discerned. Indeed, it coincides with the region of application of the A. N. Krylov method for the matrix M , proceeding from the vector R' .

As an example we shall again take the Leverrier matrix. We carry through the computation of the coefficients of the characteristic polynomial in accordance with the scheme described (see Table II). At the outset we compute matrix (1), arranging its elements in the first four rows. We next perform the elimination as was shown in § 12. The last row gives the sought values of the coefficients, which as regards accuracy almost coincide with the values computed by the method of A. N. Krylov. The last column, as usual, is the check column.

The number of operations necessary to determine the coefficients of the characteristic polynomial by the Samuelson method is somewhat less than by the A. N. Krylov method, for the formation of matrix (1) requires $n(n-1)^2$ multiplications, and the process of elimination in the Samuelson scheme requires just as many operations as does the solution of the system in the A. N. Krylov method. In connection with this we comment that in Wayland's article [1] the reckoning of the number of operations for the Samuelson method is incorrectly done.

§ 24. THE METHOD OF A. M. DANILEVSKY

An elegant and very efficient method of computing the coefficients of the characteristic polynomial has been proposed by A. M. Danilevsky [1]. The gist of his method consists in an initial reduc-

tion of the secular determinant to the form known as the Frobenius normal form:

$$D(\lambda) = \begin{vmatrix} p_1 - \lambda & p_2 & p_3 & \dots & p_n \\ 1 & -\lambda & 0 & \dots & 0 \\ 0 & 1 & -\lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\lambda \end{vmatrix},$$

the expansion of which in powers of λ demands no labor, since

$$D(\lambda) = (-1)^n [\lambda^n - p_1 \lambda^{n-1} - \dots - p_n].$$

On the strength of the fact that similar matrices have identical characteristic polynomials, in order to attain the aim set, it is sufficient to reduce the given matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

to the form

$$P = \begin{pmatrix} p_1 & p_2 & \dots & p_{n-1} & p_n \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

by means of a similarity transformation.

We shall show that we can find $(n-1)$ similarity transformations, the successive performance of which will realize the desired transition from matrix A to matrix P , if this is possible.

Let us examine the beginning of the process.

We must carry the row $a_{n1} \ a_{n2} \ \dots \ a_{n, n-1} \ a_{nn}$ into the row $0 \ 0 \ \dots \ 1 \ 0$. Let us assume that $a_{n, n-1} \neq 0$. Divide all elements of the row by $a_{n, n-1}$; then subtract the $(n-1)$ th column, multiplied by $a_{n1}, a_{n2}, \dots, a_{nn}$ respectively, from all the rest of the columns. The desired transformation will thereby obviously have been effected.

The operations we have indicated will thus be elementary transformations on columns, and, as has been shown in § 1 Paragraph 10, will reduce to a postmultiplication of the matrix A by a matrix M_{n-1} which is not hard to jot down, to wit:

$$M_{n-1} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{a_{n1}}{a_{n,n-1}} & -\frac{a_{n2}}{a_{n,n-1}} & \dots & \frac{1}{a_{n,n-1}} & -\frac{a_{nn}}{a_{n,n-1}} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

The constructed matrix AM_{n-1} will not be similar to the matrix A . However from it one can easily pass to a matrix that is similar to A . To accomplish this it is enough to premultiply the matrix AM_{n-1} by the matrix M_{n-1}^{-1} , which is easily computed; indeed one can directly verify that

$$M_{n-1}^{-1} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & a_{nn} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

The multiplication of the matrix AM_{n-1} on the left by the matrix M_{n-1}^{-1} obviously does not change the transformed row.

Thus the matrix

$$M_{n-1}^{-1}AM_{n-1} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1,n-1} & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2,n-1} & c_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ c_{n-1,1} & c_{n-1,2} & \dots & c_{n-1,n-1} & c_{n-1,n} \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} = C$$

has one satisfactory row, and we can continue the process another step further if $c_{n-1,n-2} \neq 0$. If all the $(n-1)$ transformations are possible, we shall have brought the matrix A into form P .

An explanation of how one is to proceed in those exceptional cases when the necessary elements vanish will be given below.

At the moment we shall show how the elements of the matrix C , the result of one similarity transformation, are computed.

We have

$$AM_{n-1} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & a_{nn} \end{pmatrix} \times \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ m_{n-1,1} & m_{n-1,2} & \dots & m_{n-1,n-1} & m_{n-1,n} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1,n-1} & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2,n-1} & b_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ b_{n-1,1} & b_{n-1,2} & \dots & b_{n-1,n-1} & b_{n-1,n} \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} = B,$$

where

$$b_{ik} = a_{ik} + a_{i,n-1}m_{n-1,k} \quad \text{for all } i \leq n-1, \quad k \neq n-1, \\ b_{i,n-1} = a_{i,n-1}m_{n-1,n-1}.$$

Here

$$m_{n-1,i} = -\frac{a_{ni}}{a_{n,n-1}} \quad \text{for } i \neq n-1$$

and

$$m_{n-1,n-1} = \frac{1}{a_{n,n-1}}.$$

We see that all the elements of the matrix $B = AM_{n-1}$, with the exception of the elements of the $(n-1)$ th column, are to be computed by formulas involving two terms.

The elements of the $(n-1)$ th column, if we recall the value of the coefficients $m_{n-1,n-1}$, are obtained as the quotients of the division of

TABLE III. Computational Scheme for the Method of A. M. Danilevsky

	I	1	2	3	4	Σ	Σ'
1		-5.509882	1.870086	0.422908	0.008814	-3.208074	
2		0.287865	-11.811654	5.711900	0.058717	-5.753172	
3		0.049099	4.308033	-12.970687	0.229326	-8.384229	
4		0.006235	0.269851	1.397369	-17.596207	-15.922752	
4'	$M_3^{-1} \psi \vec{M}_3$	-0.004462	-0.193114	0.71563059	12.592384	11.394808	
5		0.006235	1.788417	0.302646	5.334234	1.913527 ⁽⁸⁾	1.610881
6		0.269851	-12.914702	4.087610	71.985155	63.420442	59.332832
7		1.397369	6.812854	9.282220	-163.10255	-165.46494	-156.18272
8		-17.596207	0	1	0	1	
9		0.185920	6.046177	-29.461961	-208.45592	-231.68578	
9'	$M_2^{-1} \psi \vec{M}_2$	-0.030750	0.16539377	4.872825	34.47731	38.31938	
10		0.185920	0.295793	9.017289	66.99404	70.74035 ⁽⁶⁾	70.44456
11		6.046177	-2.136011	-58.84347	-373.2790	-433.5989 ⁽⁹⁰⁾	-431.4629
12		-29.461961	0	0	0	1	
13		-208.45592	0	1	0	1	
14		2.952517	-42.32167	-562.5575	-2244.455	-2846.382 ⁽⁶⁾	
14'	$M_1^{-1} \psi \vec{M}_1$	0.33869407	14.33410	190.5349	760.1836	964.0527	
15		-1.885430	-79.49874	-1051.645	-4164.768	-5297.798 ⁽⁷⁾	-5295.913
16		1	0	0	0	1	
17		0	1	0	0	1	
18		0	0	1	0	1	
19		-47.88843	-797.2789	-5349.455	-12296.55	-18491.17	

the elements of the $(n-1)$ th column of the matrix A by the element $a_{n, n-1}$.

Moreover,

$$C = M_{n-1}^{-1} A M_{n-1} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{n, n-1} & a_{nn} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \times \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1, n-1} & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2, n-1} & b_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ b_{n-1, 1} & b_{n-1, 2} & \dots & b_{n-1, n-1} & b_{n-1, n} \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1, n-1} & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2, n-1} & b_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \sum_{k=1}^n a_{nk} b_{k1} & \sum_{k=1}^n a_{nk} b_{k2} & \dots & \sum_{k=1}^n a_{nk} b_{k, n-1} & \sum_{k=1}^n a_{nk} b_{kn} \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

Thus the premultiplication by M_{n-1}^{-1} changes only the $(n-1)$ th row of the matrix B ; the elements of this row are found with one accumulation. The entire process, consisting of $(n-1)$ similarity transformations by means of the matrices $M_{n-1}, M_{n-2}, \dots, M_1$, is fitted into a convenient computational scheme, which we shall show in an example.

As the example we again take the Leverrier matrix (Table III). The Danilevsky transformation consists in the successive multiplication of the given matrix on the right by the matrices M_3, M_2 , and M_1 , and on the left by the matrices $M_3^{-1}, M_2^{-1}, M_1^{-1}$, where the right and left multiplications alternate. Each matrix of the transformation is given as one row.

Let us pass on to the description of the computational process. In rows 1, 2, 3, 4 are arrayed the elements of the given matrix and the check sums. We begin the computation with the computation of the elements defining the matrices M_3 and M_3^{-1} , viz.: in row 4' we enter the elements of the third row of the matrix M_3 :

$$m_{31} = -\frac{a_{41}}{a_{43}}; \quad m_{32} = -\frac{a_{42}}{a_{43}}; \quad m_{33} = \frac{1}{a_{43}}; \quad m_{34} = -\frac{a_{44}}{a_{43}}.$$

The elements of the third row of the matrix M_3^{-1} , equal to the elements a_{4k} ($k=1, 2, 3, 4$) we enter, for computational convenience, in a column: rows 5, 6, 7, 8 of the head-column I. The result of the multiplication by M_3 we write in the four rows 5, 6, 7, 8. The transformation M_3 brings the fourth row of the matrix into canonical form (the 8th row). The elements of the third column are now obtained by multiplying the elements of the third column of the matrix A by $m_{33} = \frac{1}{a_{43}}$. The rest of the elements are computed by the two-term formulas:

$$(1) \quad b_{ik} = a_{ik} + a_{i3}m_{3k}.$$

For the check we form the column of sums, as usual. The element of the check column located in row 4' must coincide with the sum of the elements of this row after replacing the element m_{33} by -1 . The results of the application of formulas (1) to the elements Σ are entered in the column Σ' ; adding the elements of the 3rd column to them, we obtain the check sums for the rows 5 to 8.

The transformation M_3^{-1} changes only the 7th row; we enter the result in the 9th row. Its elements are obtained as the sum of the products, by pairs, of the elements situated in column I by the corresponding elements of each column of the matrix AM_3 . With this we have concluded the transformation $M_3^{-1}AM_3$. The process is continued analogously. We remark that the elements of the 9th row are also the elements of the 2nd row of the matrix M_2^{-1} ; we copy them columnwise in column I. The elements of the second row of the matrix M_2 are entered in row 9'.

The process is easily learned and is afterwards executed without difficulty.

As the result of the computation we obtain all the coefficients at

once, since the coincidence—the check—of p_1 with the trace of the matrix is in addition an index to the accuracy of the computation of the rest of the coefficients; the results obtained are closer to the Leverrier data than are the results found by A. N. Krylov's method or those by the Samuelson method.

The number of operations necessary for computation by the A. M. Danilevsky method is substantially less than by the two other methods referred to above. The number of multiplications and divisions is equal to $(n-1)(n^2+n-1)$.

We remark that A. M. Danilevsky's method admits of a modification similar to the pivotal condensation scheme in connection with the Gauss method; it somewhat increases the accuracy of the results obtained.

Knowledge of the matrices M_1, M_2, \dots, M_{n-1} which successively effect the similarity transformations, permits us to determine the proper vectors of the matrix A .

Indeed, let λ be some proper number of matrix A and $X = (x_1, \dots, x_n)$ the proper vector corresponding to it. Let $Y = (y_1, y_2, \dots, y_n)$ be a proper vector of the matrix P . Then, as was shown in § 3 Paragraph 10,

$$X = M_{n-1}M_{n-2} \dots M_2M_1Y.$$

Now, the proper vector of the matrix P is found without trouble, since its components will be the solutions of the recurrence system:

$$\begin{aligned} (p_1 - \lambda)y_1 + p_2y_2 + \dots + p_ny_n &= 0 \\ y_1 - \lambda y_2 &= 0 \\ \dots &\dots \\ y_{n-1} - \lambda y_n &= 0. \end{aligned}$$

This gives $y_n = 1, y_{n-1} = \lambda, \dots, y_1 = \lambda^{n-1}$.

The transformation M_1 , performed upon Y , gives

$$M_1Y = \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n m_{1k}y_k \\ y_2 \\ \dots \\ y_n \end{pmatrix}.$$

$S=S^{-1}$, so that a multiplication on left and right by S is a similarity transformation. After these transformations one must proceed as usual.

If, however, all $c_{ki}=0$ for $i \leq k-1$, then the matter becomes even simpler.

To wit, in this case the matrix C has the form

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1, k-1} & c_{1, k} & \cdots & c_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{k-1, 1} & c_{k-1, 2} & \cdots & c_{k-1, k-1} & c_{k-1, k} & \cdots & c_{k-1, n} \\ 0 & 0 & \cdots & 0 & c_{kk} & \cdots & c_{kn} \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} = \begin{pmatrix} C_1 & D \\ 0 & C_2 \end{pmatrix},$$

where

$$C_1 = \begin{pmatrix} c_{11} & \cdots & c_{1, k-1} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ c_{k-1, 1} & \cdots & c_{k-1, k-1} \end{pmatrix}, \quad C_2 = \begin{pmatrix} c_{kk} & \cdots & c_{kn} \\ 1 & \cdots & 0 \\ \cdot & \cdot & \cdot \\ 0 & \cdots & 1 & 0 \end{pmatrix}$$

and consequently

$$|C - \lambda I| = |C_1 - \lambda I| \cdot |C_2 - \lambda I|.$$

The matrix C_2 already has the Frobenius canonical form and therefore $|C_2 - \lambda I|$ is computed instantly. In order to expand $|C_1 - \lambda I|$, one must apply the general process leading to the matrix P_1 .

Thus the case where the process is broken introduces only simplifications into the problem of computing the characteristic polynomial: an easily computable factor is separated from the characteristic polynomial, and the remaining factor is the characteristic polynomial of a matrix of lower order.

§ 25. LEVERRIER'S METHOD IN D. K. FADDEEV'S MODIFICATION

In this section we shall expound a method known as Leverrier's method [2], requiring a greater number of operations than any of the methods presented above, but utterly insensitive to the individual peculiarities of the matrix, in particular to "gaps" in the intermediate determinants.

Let

$$(1) \quad D_n(\lambda) = (-1)^n [\lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \cdots - p_n]$$

be the characteristic polynomial of the matrix and $\lambda_1, \lambda_2, \dots, \lambda_n$ its roots, among which some may be equal. Let us employ the symbol

$$(2) \quad \sum_{i=1}^n \lambda_i^k = s_k.$$

Then a relation is valid that is known as the Newton formula:

$$(3) \quad k p_k = s_k - p_1 s_{k-1} - \cdots - p_{k-1} s_1, \quad k=1, \dots, n.$$

If the numbers s_k are known, then by solving the recurrence system (3) we can find the coefficients p_k which we need.

We shall show how the numbers s_k are determined. We have

$$s_1 = \lambda_1 + \lambda_2 + \cdots + \lambda_n = \text{tr } A.$$

Moreover, on the strength of Paragraph 11, § 3, the characteristic numbers of the matrix A^k will be $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$. Accordingly

$$(4) \quad s_k = \lambda_1^k + \lambda_2^k + \cdots + \lambda_n^k = \text{tr } A^k.$$

Thus the process of computation reduces to the successive computation of the powers of the matrix A , then to the calculation of their traces and, finally, to the solution of the recurrence system (3). The computation of the n powers of the matrix A (of the last matrix A^n it is necessary to compute only the diagonal elements) requires a great number of operations—uniform, granted—and Leverrier's method is inordinately more laborious than the methods expounded above. Its value consists, as has already been mentioned, in its