Let B denote the collection of all the classes $C_{(r,x)}$. Then B is an algebra. The subset A' of B consisting of all the classes $C_{(1,x)}$ of B is isomorphic with A and is identical with B if, and only if, for each element x of A and each positive rational integer m, A contains an element x/m such that $m(x/m) = x$.

If $C_{(r,x)}$ is an element of B and m is a positive rational integer, then $C_{(r/m,x)}$ is an element of B and $mC_{(r/m,x)} = C_{(r,x)}$.

Multiplication in B is associative, commutative, distributive, respectively, if the same is true for A.

PRINCETON UNIVERSITY

ON THE RANK OF THE PRODUCT OF CERTAIN SQUARE MATRICES*

BY W. O. MENGE

1. *Introduction.* This paper presents several theorems which were found during an investigation conducted by the author into the structure of matrices which transform given matrices into their so-called classic and rational canonical forms.† When the elementary divisors of a given matrix are known, these theorems completely determine the rank of a product of matrices of the form

$$\prod_{i=1}^{\omega}(A - \lambda_i I)^{k_i}.$$

An interesting proof of the Hamilton-Cayley theorem and a determination of the equation of minimum degree satisfied by a matrix are obtained from this point of view.

2. *Invariant Factors.* Consider the square matrix $A = (a_{ij})$ of order n with constant elements. If the n-rowed identity matrix be denoted by I, the characteristic matrix $(A - \lambda I)$ is defined as the matrix obtained by subtracting the variable λ from each principal diagonal element of A. The determinant, $D(\lambda)$, of the characteristic matrix $(A - \lambda I)$ is called the characteristic de-

† Dickson, *Modern Algebraic Theory*, Chap. 5.

terminant of A. In order to secure complete generality in the subsequent discussion, the characteristic equation $D(\lambda) = 0$ is assumed to have the distinct roots $\lambda_1, \lambda_2, \cdots, \lambda_\omega$ of multiplicities $n^{(1)}, n^{(2)}, \cdots, n^{(\omega)}$, respectively. Hence

$$D(\lambda) = (\lambda - \lambda_1)^{n^{(1)}}(\lambda - \lambda_2)^{n^{(2)}} \cdots (\lambda - \lambda_\omega)^{n^{(\omega)}},$$

where $\sum_{i=1}^{\omega} n^{(i)} = n$.

Designate by $G_j(\lambda)$ the highest common factor of all jth minors (of order $n-j$) of the characteristic determinant $D(\lambda)$. These common factors are chosen so that the coefficient of the highest power of λ is unity. Thus one can write

$$G_j(\lambda) = (\lambda - \lambda_1)^{m_j^{(1)}}(\lambda - \lambda_2)^{m_j^{(2)}} \cdots (\lambda - \lambda_\omega)^{m_j^{(\omega)}}, \quad (j = 1, 2, \cdots, n).$$

The invariant factors of the matrix A are defined as

$$D_j(\lambda) = G_{j-1}(\lambda)/G_j(\lambda) = (\lambda - \lambda_1)^{n_j^{(1)}}(\lambda - \lambda_2)^{n_j^{(2)}} \cdots (\lambda - \lambda_\omega)^{n_j^{(\omega)}}$$

$$= \lambda^{n_j} - d_{1j} - d_{2j}\lambda - \cdots - d_{n_j,j}\lambda^{n_j-1}, \quad (j = 1, 2, \cdots, n),$$

where

$$G_0(\lambda) = D(\lambda), \qquad n_j^{(i)} = m_{j-1}^{(i)} - m_j^{(i)}, \qquad n_j = \sum_{i=1}^{\omega} n_j^{(i)},$$

$$(i = 1, 2, \cdots, \omega; j = 1, 2, \cdots, n).$$

3. *Classic Canonical Form.* Define the integer σ, so that

$$D_j(\lambda) \neq 1 \text{ for } j \leq \sigma; \quad D_j(\lambda) = 1 \text{ for } j > \sigma.$$

The classic canonical form for the n-rowed square matrix A is defined as

$$P = \begin{vmatrix} N_{11} & 0 & 0 & \cdots 0 \\ 0 & N_{21} & 0 & \cdots 0 \\ 0 & 0 & N_{31} & \cdots 0 \\ \cdot & \cdot & \cdot & \cdot \cdot \cdot \cdot \\ 0 & 0 & 0 & \cdots N_{\omega\sigma} \end{vmatrix},$$

where N_{ij} is an $n_j^{(i)}$-rowed square matrix of the form

$$N_{ij} = \begin{vmatrix} \lambda_i & 0 & 0 & \cdots 0 \\ 1 & \lambda_i & 0 & \cdots 0 \\ 0 & 1 & \lambda_i & \cdots 0 \\ \cdot & \cdot & \cdot & \cdot \cdot \cdot \\ 0 & 0 & 0 & \cdots \lambda_i \end{vmatrix}, \qquad \begin{array}{l} (i = 1, 2, \cdots, \omega; \\ j = 1, 2, \cdots, \sigma). \end{array}$$

Each of the matrices N_{ij} is associated with a particular elementary divisor. From the above definition it can be easily verified that the invariant factors and elementary divisors of the matrix P are identical with the invariant factors and elementary divisors of the matrix A.

4. *Ranks.* It has been shown* that a necessary and sufficient condition that a non-singular matrix B exist such that

$$A = B \cdot P \cdot B^{-1}$$

is that the matrices A and P have the same invariant factors, or if we prefer, the same elementary divisors. It follows immediately that a non-singular matrix B exists such that

(1) $A - \lambda I = B \cdot (P - \lambda I) \cdot B^{-1}$.

Since equation (1) is an identity in the variable λ, one can write

(2) $\prod_{i=1}^{\omega}(A - \lambda_i I)^{k_i} = B^m \cdot \prod_{i=1}^{\omega}(P - \lambda_i I)^{k_i} \cdot (B^{-1})^m$

$$= B^m \cdot \prod_{i=1}^{\omega}(P - \lambda_i I)^{k_i} \cdot (B^m)^{-1},$$

where $m = \sum_{i=1}^{\omega} k_i$. Since the matrix B is non-singular it follows that B^m is also non-singular. Hence we may state the following lemma.

LEMMA 1. *The rank of the product*

$$\prod_{i=1}^{\omega}(A - \lambda_i I)^{k_i}$$

is identical with the rank of the product

$$\prod_{i=1}^{\omega}(P - \lambda_i I)^{k_i}.$$

It will be seen that the rank of any given product of the latter type can be determined by the invariant factors and elementary divisors of the matrix A.

Using the notation of Kronecker, define a matrix of order q, $(\delta_{r,s+1}) = (c_{rs})$, where $c_{rs} = 1$ for $r = s+1$, $c_{rs} = 0$ for $r \neq s+1$. It is easy to verify that $(\delta_{r,s+1})^k = (\delta_{r,s+k})$, and to prove the following result.

---

* Bôcher, *Introduction to Higher Algebra*, pp. 279–283.

LEMMA 2. *The matrix* $(\delta_{r,s+k})$ *is of rank* $q-k$ *if* $q-k \geqq 0$ *and of zero rank if* $q-k < 0$.

In forming the product $(P-\lambda_i I)^{k_i}$ from its linear factors it is apparent from the special nature of each factor that the determinants $|N_{h_j}-\lambda_i I|$ associated with each elementary divisor multiply each other independently of the determinants associated with other elementary divisors. The product is of the same form as each factor, and the determinant associated with each elementary divisor in the product is made up of the product of the corresponding determinants in the several factors. The determinants $|N_{ij}-\lambda_i I|$ are of the form $(\delta_{r,s+1})$ with $q=n_j^{(i)}$, and the determinants $|N_{hj}-\lambda_i I|$ with $h \neq i$ are non-singular.

Define $r_i^{(l)}$ as the number of elementary divisors of A divisible by $(\lambda-\lambda_i)^l$. Then $r_i^{(l)}$ will also be the number of elementary divisors of P which contain the factor $(\lambda-\lambda_i)^l$. Upon applying Lemma 2 for the multiplication of matrices of this form it is seen that the rank of the determinant associated with the elementary divisor $(\lambda-\lambda_i)^{n_j^{(i)}}$ in the product $(P-\lambda_i I)^{k_i}$ will be $n_j^{(i)}-k_i$ if $n_j^{(i)}-k_i \geqq 0$ and will be zero if $n_j^{(i)}-k_i < 0$. The determinants associated with elementary divisors which do not contain the root λ_i will be non-singular. Thus the rank of the matrix of the product $(P-\lambda_i I)^{k_i}$ is seen to be

$$n - \sum_{l=1}^{k_i} r_i^{(l)}.$$

In the product $(P-\lambda_i I)^{k_i}(P-\lambda_j I)^{k_j}$, $j \neq i$, the determinants associated with elementary divisors involving either of the roots λ_i, λ_j are singular and of the form $(\delta_{r,s+1})$, while those associated with elementary divisors involving other roots are non-singular. Hence the rank of the product $(P-\lambda_i I)^{k_i}(P-\lambda_j I)^{k_j}$ is

$$n - \sum_{l=1}^{k_i} r_i^{(l)} - \sum_{l=1}^{k_j} r_j^{(l)}.$$

In general we find that the rank of the product

$$\prod_{i=1}^{\omega}(P - \lambda_i I)^{k_i}$$

is

$$n - \sum_{i=1}^{\omega} \sum_{l=1}^{k_i} r_i^{(l)}.$$

Upon applying Lemma 1 to the result just obtained we have the following theorem.

THEOREM 1. *The rank of the product*

$$\prod_{i=1}^{\omega} (A - \lambda_i I)^{k_i}$$

is

(i)
$$n - \sum_{i=1}^{\omega} \sum_{l=1}^{k_i} r_i^{(l)}.$$

This result may also be written in the form

(ii)
$$n - \sum_{i=1}^{\omega} \left[k_i r_i^{(k_i)} + \sum_{n_h^{(i)} < k_i} n_h^{(i)} \right].$$

Define $D_j(A)$ as the matrix obtained by replacing the variable λ by the matrix A in the jth invariant factor $D_j(\lambda)$. Thus

$$D_j(A) = (A - \lambda_1 I)^{n_j^{(1)}} (A - \lambda_2 I)^{n_j^{(2)}} \cdots (A - \lambda_\omega I)^{n_j^{(\omega)}}.$$

Any matrix $f(A)$, where $f(\lambda)$ is of equal or lower degree in λ than $D_j(\lambda)$ can be obtained from $D_j(A)$ in one of the following ways:

(a) by replacing one or more factors, say $(A - \lambda_1 I)$, by another, say $(A - \lambda_2 I)$;

(b) by omitting one or more factors, say $(A - \lambda_1 I)$, or by replacing them by non-singular matrices.

Under Case (a), the rank R of the altered matrix is, according to Theorem 1,

$$R = n - \sum_{l=1}^{n_j^{(1)}-1} r_1^{(l)} - \sum_{l=1}^{n_j^{(2)}+1} r_2^{(l)} - \sum_{i=3}^{\omega} \sum_{l=1}^{n_j^{(i)}} r_i^{(l)}.$$

Similarily the rank R_j of the matrix $D_j(A)$ is

$$R_j = n - \sum_{i=1}^{\omega} \sum_{l=1}^{n_j^{(i)}} r_i^{(l)},$$

and the difference in rank is

$$R - R_j = r_1^{(n_j^{(1)})} - r_2^{(n_j^{(2)}+1)}.$$

It is apparent, however, from the definition of $r_i^{(l)}$ and the prop-

erty of each invariant factor being divisible by each succeeding invariant factor, that

$$r_1^{(n_j^{(1)})} \geq j, \quad \text{while} \quad r_2^{(n_j^{(2)}+1)} < j.$$

From this it follows immediately that $R \geq R_j + 1$ and hence the replacement in Case (a) increases the rank by at least unity.

Under Case (b) by Theorem I the rank R of the altered matrix is determined as

$$R = n - \sum_{l=1}^{(n_j^{(1)}-1)} r_1^{(l)} - \sum_{i=2}^{\omega} \sum_{l=1}^{n_j^{(i)}} r_i^{(l)},$$

from which one concludes

$$R - R_j = r_1^{(n_j^{(1)})}.$$

Hence the rank R is larger than the rank R_j of the invariant factor matrix $D_j(A)$. Thus one has established the following fact.

THEOREM 2. *The rank of any invariant factor matrix $D_j(A)$ is the minimum for all matrices $f(A)$, where $f(\lambda)$ is of the same or lower degree in λ than $D_j(\lambda)$.*

By setting $k_i = n_j^{(i)}$, $(i = 1, 2, \cdots, \omega)$, the result (ii) expressed in Theorem I can be applied to determine the rank R_j of $D_j(A)$. Here we find

$$(3) \qquad R_j = n - \sum_{i=1}^{\omega} \left[n_j^{(i)} \cdot r_i^{(n_j^{(i)})} + \sum_{n_h^{(i)} < n_j^{(i)}} n_h^{(i)} \right],$$

the summation inside the bracket being with respect to h. Setting $s_j^{(i)} = r_i^{(n_j^{(i)})} - j$ in equation (3), we obtain

$$R_j = n - \sum_{i=1}^{\omega} \left[n_j^{(i)}(j + s_j^{(i)}) + \sum_{n_h^{(i)} < n_j^{(i)}} n_h^{(i)} \right],$$

and

$$(4) \qquad R_j = n - j \cdot n_j - \sum_{i=1}^{\omega} \left[s_j^{(i)} n_j^{(i)} + \sum_{n_h^{(i)} < n_j^{(i)}} n_h^{(i)} \right],$$

since

$$\sum_{i=1}^{\omega} n_j^{(i)} = n_j.$$

Equation (4) may be written in the form

$$R_j = n - j \cdot n_j - \sum_{i=1}^{\omega} \sum_{h=j+1}^{\sigma} n_h^{(i)},$$

or

$$(5) \qquad R_j = n - j \cdot n_j - \sum_{h=j+1}^{\sigma} n_h.$$

By virtue of the equation

$$\sum_{h=1}^{j} n_h = n - \sum_{h=j+1}^{\sigma} n_h,$$

the expression (5) may be written in the form

$$R_j = \sum_{h=1}^{j} n_h - j \cdot n_j = \sum_{h=1}^{j}(n_h - n_j).$$

THEOREM 3. *The rank R_j of $D_j(A)$ is*

$$\sum_{h=1}^{j}(n_h - n_j) = \sum_{h=1}^{j} n_h - j \cdot n_j = n - j \cdot n_j - \sum_{h=j+1}^{\sigma} n_h.$$

It should be noted that two well known properties of the matrix A follow directly from Theorem 3, namely, *Every square matrix satisfies its characteristic equation* and *The equation of minimum degree satisfied by the matrix A is the first invariant factor equated to zero.*

THEOREM 4. *In progressing from $D_{j+1}(A)$ to $D_j(A)$ by multiplying by successive linear factors, the rank decreases j for each linear factor.*

THEOREM 5. *The difference in the ranks of two consecutive invariant factor matrices, $D_{j+1}(A)$ and $D_j(A)$, is $j(n_j - n_{j+1})$.*

THE UNIVERSITY OF MICHIGAN