ON THE REDUCTION OF A MATRIX TO ITS RATIONAL CANONICAL FORM*

BY M. H. INGRAHAM

Two square $n\times n$ matrices A and B with elements in a field F are similar in F if there exists a non-singular $n\times n$ matrix with elements in F such that $S^{-1}AS=B$. In the study of similarity canonical forms play a fundamental role. The classical canonical form for A is one in which the elementary divisors of $A-\lambda I$ are brought into prominence. In 1926, Dickson published in his *Modern Algebraic Theories* a rational discussion of the problem of similarity in which a rational canonical form based on the invariant factors of $(A-\lambda I)$ was used. Other discussions by Lattés, Krull, Kowalewski, and Menge have been published.

The following seems to be, from the algebraic standpoint, a somewhat more direct discussion than others known to the author. Moreover, in arriving at the well known rational canonical form for a matrix, certain lemmas of interest are developed.

Throughout we consider all elements of matrices and vectors and coefficients of polynomials that enter the discussion to be in a field F. All points of interest are met with if the elements involved are rational.

Consider an $n\times n$ matrix A. If the vectors $\xi_1, \xi_2, \cdots, \xi_p$ are $n\times 1$ matrices, we define $L(\xi_1, \xi_2, \cdots, \xi_p)$ to be the linear set consisting of all vectors of the form $\sum_1^p g_i(A)\xi_i$, where the g's are polynomials.

If L is such a linear set and if g is a polynomial and η an $n\times 1$ vector, we say that $g(A)\eta\equiv 0 \bmod L$, where 0 stands for the zero vector if $g(A)\eta$ is in L.

If $g_1(A)\eta\equiv 0 \bmod L$, and $g_2(A)\eta\equiv 0 \bmod L$, then for every pair of polynomials p_1, p_2,

$$\{p_1(A)g_1(A) + p_2(A)g_2(A)\}\eta \equiv 0 \bmod L.$$

Since the greatest common divisor of g_1 and g_2 is expressible in the form $p_1g_1+p_2g_2$, one can prove (as in the proof for the existence of a minimum equation for a matrix) that for each

* Presented to the Society, November 25, 1932.

$n\times 1$ vector η and linear set L there exists a polynomial g of lowest degree and leading coefficient unity such that $g(A)\eta\equiv 0$ mod L. In fact any other polynomial satisfying this congruence will be divisible by g. The polynomial g is called the minimum function of A relative to (η, L).

If g_1 and g_2 are polynomials, then there exist polynomials f_1, f_2, h_1, and h_2 such that $g_1=f_1h_1$, $g_2=f_2h_2$, where f_1 and f_2 are relatively prime and f_1f_2 is the least common multiple of g_1 and g_2. This is readily seen, for if the factorizations of g_1 and g_2 into powers of distinct irreducible factors are given by

$$g_1 = p_1^{m_1}p_2^{m_2}\cdots p_s^{m_s}q_1^{n_1}q_2^{n_2}\cdots q_t^{n_t},\ g_2 = p_1^{v_1}p_2^{v_2}\cdots p_s^{v_s}q_1^{u_1}q_2^{u_2}\cdots q_t^{u_t},$$

where $0<m_i\geqq v_i\geqq 0$ and $0\leqq n_i\leqq u_i>0$, then $f_1=p_1^{m_1}p_2^{m_2}\cdots p_s^{m_s}$ and $f_2=q_1^{u_1}q_2^{u_2}\cdots q_t^{u_t}$ are effective. If g_1 is the minimum function of A relative to (η_1, L) and g_2 is the minimum function of A relative to (η_2, L) and if $\eta=h_1(A)\eta_1+h_2(A)\eta_2$, then g, the minimum function of A relative to (η, L), is f_1f_2, the least common multiple of g_1 and g_2. For we have

$$g(A)\eta = g(A)h_1(A)\eta_1 + g(A)h_2(A)\eta_2 \equiv 0 \bmod L,$$

and hence

$$f_1(A)g(A)h_1(A)\eta_1 + f_1(A)g(A)h_2(A)\eta_2 \equiv 0 \bmod L.$$

But $f_1(A)h_1(A)\eta_1\equiv 0 \bmod L$, and thus $f_1(A)g(A)h_2(A)\eta_2\equiv 0 \bmod L$, and f_1gh_2 is a multiple of g_2 and hence g a multiple of f_2. Similarly, g is a multiple of f_1 and since $f_1(A)f_2(A)\eta\equiv 0$ mod L, $g=f_1f_2$. Hence if the minimum function of A relative to (η_1, L) is g_1 and if there exists another vector η_2 for which the minimum function of A relative to (η_2, L) is g_2, where g_2 does not divide g_1, there exists a third vector η_3 such that the minimum function of A relative to (η_3, L) is of higher degree than either g_1 or g_2. As the degree of any such minimum function can not exceed n, there exists a vector η such that g, the minimum function of A relative to (η, L), is divisible by the minimum function of A relative to (η_1, L), where η_1 is any arbitrary vector. Such an η is said to be maximal relative to (A, L) and g is called the minimum function of A relative to L. This process of finding the maximal η corresponds to the finding of the leader of a chain of maximal length at various stages in Dickson's discussion. The present method, however, gives a mode of construction for this leader.

It is readily seen that if L_1 is contained in L_2, then for any vector η the minimum function of A relative to any vector (η, L_2) is a divisor of the minimum function of A relative to (η, L_1), and that the minimum function of A relative to L_2 is a divisor of the minimum function of A relative to L_1.

Consider a set of vectors $\xi_1, \xi_2, \cdots, \xi_p$. Let $L_0=0$, $L_1=L(\xi_1)$, $L_j=L(\xi_1, \cdots, \xi_j)$. Let ξ_1 be maximal relative to (A, L_0), ξ_2 maximal relative to (A, L_1), and, in general, ξ_j maximal relative to (A, L_{j-1}). If L_p is not the complete vector space, there exists a vector η not in L_p which is maximal relative to (A, L_p). Let g_i be the minimum function of A relative to L_{i-1}. From the preceding paragraph g_{i-1} is divisible by g_i for each i; hence a set of polynomials k_i exists, such that $g_i=k_ig_{p+1}$, $(i<p+1)$. Since $g_{p+1}(A)\eta\equiv 0 \bmod L_p$, there exist a set of polynomials of f_i such that

$$g_{p+1}(A)\eta = \sum_1^p f_i(A)\xi_i.$$

Moreover, each f_i is divisible by g_{p+1}, for if this were not the case there would be a last number l less than or equal to p for which f_l is not divisible by g_{p+1}. But if j is greater than l, k_lf_j is a multiple of g_l since f_j is assumed to be a multiple of g_{p+1}; hence $k_l(A)f_i(A)\xi_i\equiv 0 \bmod L_{l-1}$ whenever $i\neq l$. Since $k_l(A)g_{p+1}(A)\eta = g_l(A)\eta\equiv 0 \bmod L_{l-1}$, it follows from the last remark that

$$k_l(A)f_l(A)\xi_l \equiv 0 \bmod L_{l-1},$$

but since ξ_l is maximal relative to (A, L_{l-1}), k_lf_l must be a multiple of $g_l=k_lg_{p+1}$ and hence f_l must be a multiple of g_{p+1} in contradiction to our hypothesis. Hence, if $f_i=r_ig_{p+1}$ and if we let

$$\xi_{p+1} = \eta - \sum_1^p r_i(A)\xi_i,$$

it follows that $g_{p+1}(A)\xi_{p+1}=0$ and hence ξ_{p+1}, which is maximal relative to L_p, is such that A has g_{p+1} for its minimum function relative to (ξ_{p+1}, L_0) as well as relative to (ξ_{p+1}, L_p). Hence we may find a set of vectors $\xi_1, \cdots, \xi_t$ with the properties:

(1) Each ξ_i is maximal relative to $L(\xi_1, \cdots, \xi_{i-1})$.

(2) If g_i is the minimum function of A relative to $(\xi_i, L(\xi_1, \cdots, \xi_{i-1}))$, then $g_i(A)\xi_i=0$.

(3) $L(\xi_1, \cdots, \xi_t)$ is the total vector space.

(4) Each ξ_i is different from the zero vector.

Let $g_i(\lambda) = -\sum_1^{n_i-1} b_{ij}\lambda^j + \lambda^{n_i}$. Consider the $n\times n$ matrix, S, whose columns are the linearly independent vectors $\xi_1, A\xi_1, \cdots, A^{n_1-1}\xi_1, \xi_2, A\xi_2, \cdots, A^{n_2-1}\xi_2, \cdots, A^{n_t-1}\xi_t$. If the rows of S^{-1} are the $1\times n$ vectors $\zeta_1', \cdots, \zeta_n'$, these vectors form with the rows of the transpose of S a biorthogonal system. The columns of AS are

$$A\xi_1, A^2\xi_1, \cdots, A^{n_1-1}\xi_1, \sum_0^{n_1-1} b_{ij}A^j\xi_1, A\xi_2, \cdots, \sum_0^{n_t-1} b_{tj}A^j\xi_t,$$

and

$$\zeta_1'A\xi_1 = 0, \zeta_1'A^2\xi_1 = 0, \cdots, \zeta_1'\sum_0^{n_1-1} b_{ij}A^j\xi_1 = b_{10}, \zeta_1'A\xi_2 = 0, \cdots,$$
$$\zeta_2'A\xi_1 = 1, \zeta_2'A^2\xi_1 = 0, \cdots, \zeta_2'\sum_0^{n_1-1} b_{ij}A^j\xi_1 = b_{11}, \zeta_2'A\xi_2 = 0, \cdots,$$

which are the elements of $S^{-1}AS$, and hence $S^{-1}AS$ is 0 except for $n_i\times n_i$ blocks along the main diagonal of the form

$$\begin{pmatrix} 0 & 0 & 0\cdots 0 & b_{i0} \\ 1 & 0 & 0\cdots 0 & b_{i1} \\ 0 & 1 & 0\cdots 0 & b_{i2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0\cdots 1 & b_{in_i-1} \end{pmatrix}$$

whose characteristic determinant is $g_i(\lambda)$, where each $g_i(\lambda)$ divides $g_{i-1}(\lambda)$, and hence $S^{-1}AS$ is the transpose of the rational canonical form given by Dickson.

The uniqueness of this canonical form may be established as usual through the invariance of the greatest common divisors of the determinants of the $(n-r)$-rowed minors of $A-\lambda I$, or as follows. Since $g(S^{-1}AS) = S^{-1}g(A)S$, the rank of a polynomial in A and the rank of the same polynomial in any matrix similar to A are identical. Consider two distinct $n\times n$ matrices A_1 and A_2 in canonical form with the characteristic functions of the ith blocks g_{1i} and g_{2i}, respectively; if $g_{1i} = g_{2i}$ for $i<k$ and $g_{1k}\neq g_{2k}$; then if g_{1k} is not divisible by g_{2k}, $g_{2k}(A_2)$ will have rank less than $g_{2k}(A_1)$.

THE UNIVERSITY OF WISCONSIN