A rational canonical form for matrix fields*

by

J. T. B. Beard, Jr., (Arlington, Tex.)

1. Introduction and notation. Let $F$ be an arbitrary field and let $(F)^n$ denote the algebra of all $n \times n$ matrices over $F$ under normal matrix addition and multiplication. The primary purpose of this paper is to examine a rational canonical form (R.C.F.) for matrix fields over $F$. This R.C.F. is in general not unique and the obvious questions remain open. In the final section we consider a relationship between matrix roots of prime polynomials over GF(q).

In certain instances we have a technique for extending matrix fields within $(F)^n$ ([1], Theorems 9, 10, [2], Theorems 12, 13). The R.C.F. defined in § 2 is principally motivated by an unsuccessful attempt to improve and generalize that technique. In particular we ask: given a subfield $M$ of $(F)^n$ with $M$ containing a matrix in rational canonical form (r.c.f.) over $F$, can we extend $M$ non-trivially by adjoining a matrix $A \in F$ where $A$ is in r.c.f. over $F$? The negative answer raises a more general question which we consider in § 3.

Our notation and terminology is that of [1], [2] and briefly is as follows. If a matrix $A \in (F)^n$ is the matrix direct sum of $k$ companion matrices over $F$, we call $A$ a $k$-matrix and follow the convention that the coefficients of a monic polynomial $f(x) \in F[x]$ determine the last row of its companion matrix $C(f(x))$. It is well known that if $g(x) = a_n x^{n-1} + \ldots + a_1 x + a_0 \in F[x]$ and $C(f(x)) \in (F)^n$, then the first row of the matrix $g(C(f(x)))$ is given by the vector $(a_n, \ldots, a_0)$. By the r.c.f. over $F$ of a matrix $A \in (F)^n$, we mean the matrix diag $C(f_1(x)), \ldots, C(f_k(x))$, where the polynomials $f_i(x)$ are the non-trivial similarity invariants of $A$ over $F$ and $\deg f_i(x) < \deg f_{i+1}(x)$ for $1 \leq i < k$. Finally, we denote the set of all scalar matrices in $(F)^n$ by $S_n(F)$ and the set of all subfields of $(F)^n$ by $S_n$.

2. A rational canonical form. We remember from [1] that if $M \in \mathcal{F}_n$ has rank $r$, then $M$ is similar over $F$ to a matrix field $M'$ in which each matrix has the form $\text{diag} \{ O_{n-r}, A' \}$, and $A' \in (F)^r$, has rank $r$ if and only

* Portions of this paper are contained in the author's doctoral dissertation, directed by Professor Robert M. McConnell.
if the corresponding matrix \( A \in M \) is non-zero. We call \( M' \) a normal form for \( M \) over \( F \) and let \( \pi, M' \) denote the obvious projective image of \( M' \) in \( F' \).

**Definition 1.** Let \( M \in F_n \) have rank \( r \) and let \( M' \) be a normal form for \( M \) over \( F \). Then \( M' \) is called a rational canonical form (R.C.F.) for \( M \) over \( F \) if and only if \( \pi, M' \) contains a non-scalar matrix in r.c.f. over \( F \) whenever \( \pi, M' \) contains a non-scalar matrix.

Clearly, each \( M \in F_n \) having rank \( r \) is in normal form; \( M' \in F_n \) is its unique R.C.F. whenever \( \pi, M' \) contains only scalar matrices; and each \( M \in F_n \) has a R.C.F. over \( F \). While it is easy to verify that a R.C.F. is not necessarily unique, we are able to obtain the following theorem.

**Theorem 1.** Let \( F \) be an arbitrary field, and let \( M \in F_n \) have rank \( r \). If \( M' \) is any R.C.F. for \( M \) over \( F \), then \( \pi, M' \) contains at most one non-scalar matrix in r.c.f. over \( F \).

The above result follows immediately from Theorem 2.

**Theorem 2.** Let \( F \) be an arbitrary field, and let \( M \in F_n \) have rank \( r \). Then at most one non-scalar matrix in \( M \) is in r.c.f. over \( F \).

**Proof.** Suppose \( M \) contains a non-scalar matrix \( A \) in r.c.f. over \( F \), say

\[
A = \text{diag}(A_1, \ldots, A_k),
\]

where the companion matrix \( A_i \) has order \( m_i \) for \( 1 \leq i \leq k \). If \( A' \in M \) is also a non-scalar matrix in r.c.f. over \( F \), we let

\[
A' = \text{diag}(A'_1, \ldots, A'_l),
\]

where the companion matrix \( A'_i \) has order \( n_i \) for \( 1 \leq i \leq l \). Since neither \( A \) nor \( A' \) are scalar matrices, then \( k, l < n \). Let \( \pi \) and \( \pi' \) denote the ordered partitions of \( n \) as defined by \((m_1, \ldots, m_k)\) and \((n_1, \ldots, n_l)\) respectively. Let \( m = \max(m_k, n_l) \). We can assume w.l.o.g. that \( m \) belongs to the partition \( \pi \). Partition both \( A \) and \( A' \) into block matrices, say \( A = [B_{ij}] \) and \( A' = [B'_{ij}] \), where \( B_{ij} \) and \( B'_{ij} \) both have dimensions \( m_i \times m_j \) for \( 1 \leq i, j \leq k \), as determined by the partition \( \pi \). Then \( A_k = B_{kk} \) is non-derogatory. Since \( M \) is a field, \( A \) and \( A' \) commute. We conclude that \( B_{kk} \) commutes with \( B_{kk}' \), and hence that \( B_{kk}' = g(B_{kk}) \) for some \( g(x) \in F[x] \) with \( \deg g(x) < m \). Since \( B_{kk} \) is a companion matrix it follows that \( g(x) = a \) for some \( a \in F \) or else \( g(x) = x \), due to the form of the first row of \( B_{kk} \). If \( g(x) = a \) then \( A_i = [a] \), and hence \( A' = aI_k \) by the divisibility properties of the similarity invariants of \( A' \). This is a contradiction, hence \( g(x) = x \) and \( A_k = A_k' \). Thus \( A = A' \), for otherwise \( A - A' \) is non-zero and has rank less than \( n \).

In summary, we have

**Theorem 3.** Let \( F \) be an arbitrary field, and let \( M \in F_n \) have rank \( r \). Then \( M \) has a R.C.F. over \( F \). If \( M' \in F_n \) is any R.C.F. for \( M \) over \( F \) and \( K \)

is any extension field over \( F \), then \( M' \) is a R.C.F. for \( M \) over \( K \). Furthermore, \( \pi, M' \) contains at most one non-scalar matrix in r.c.f. over \( F \).

**3. k-matrices in matrix fields.** The proof technique used in Theorem 2 does not appear to be particularly fragile, so we question the uniqueness of \( k \)-matrices in matrix fields. The answer is negative as shown by this example.

**Example.** Let \( F = GF(64) \) so that \( F \) has prime subfield \( GF(2) \) and proper subfields \( GF(4) \) and \( GF(8) \). Let \( f(x) = x^2 + x + 1 \), so that \( f(x) \) is prime in \( GF[2][x] \) and splits over \( GF(4) \). Choose \( a \in GF(4) \) as a root of \( f(x) \). Let \( C_1 = C_{f(x)} \), \( A_1 = aI_2 \) where \( I_2 \) is the identity of \( (F)_2 \), and \( A = \text{diag}(C_1, A_1) \). Then \( S_{f(x)}[A, B] \in F \).

Now let \( g(x) = x^2 + x + 1 \), so that \( g(x) \) is prime in \( GF[2][x] \), and splits over \( GF(4) \), and choose a root \( a \) of \( g(x) \) in \( GF(8) \). Let \( A_2 = C_{g(x)} \), \( C_2 = aI_4 \), and \( B = \text{diag}(C_2, A_2) \). It follows that \( S_{f(x)}[A, B] \in F \), and contains the non-scalar \( 4 \)-matrix \( A_1 \) and also the non-scalar \( 3 \)-matrix \( B \).

In the other direction, it is easy to construct matrix fields in which the zero matrix is the only \( k \)-matrix.

We do gain the desired uniqueness in certain cases and are reminded of the question in [2] concerning the set of scalar matrices contained in a matrix field. We remember that if \( T \) is a subset of \( (F)_n \), then the entry field of \( T \) is the smallest subfield \( F' \) of \( F \) such that \( T \) is contained in \( (F')_n \). The method of Theorem 2 yields the following result.

**Theorem 4.** Let \( F \) be an arbitrary field. Let \( M \in F_n \) be in normal form having rank \( r \) and entry field \( F' \), and suppose \( \pi, M \) contains \( S_{f(x)}[A, B] \). If \( M \) contains a \( k \)-matrix \( A \) and an \( l \)-matrix \( A' \) with \( k, l < n \), then \( A = A' \).

**4. Other results.** In this section we sharpen and extend the following result.

**Theorem 5.** Let \( F = GF(p) \). Let \( A \in (F)_n \) have characteristic polynomial \( f(x) \) and minimal polynomial \( f(x) \) which is prime in \( F[x] \). Then for each positive divisor \( m \) such that \( m \mid n \) there exists a polynomial \( g(x) \in E[x] \) of degree \( r < n/m \) and a prime polynomial \( h(x) \in F[x] \) of degree \( m \) such that \( g(A) \) has characteristic polynomial \( h^{m/n}(x) \) and minimal polynomial \( h(x) \).

The above theorem will follow easily from Theorem 2 in [1]. We remember that if \( F = GF(q) \) is the Galois field of order \( q \) and \( A \in (F)_n \) has minimal polynomial of degree \( m \) over \( F \), then the ring extension \( S_A(F)[A] \) of \( S_A(F) \) by \( A \) has order \( q^m \) and is given by

\[
S_A(F)[A] = \{ g(A); \; g(x) \in F[x], \; \deg g(x) < m \}.
\]

We restate Theorem 5 equivalently but in simpler form.

**Theorem 6.** Let \( F = GF(p) \). Suppose \( A \in (F)_n \) has minimal polynomial \( f(x) \) which is prime in \( F[x] \) and has degree \( s \). Then for each positive
divisor $m$ of $s$, there exists a polynomial $g(x) \in F[x]$ of degree $r < s$ and a prime polynomial $h(x) \in F[x]$ of degree $m$, such that $g(A)$ has minimal polynomial $h(x)$.

Proof. The ring $S_\infty(F)[A]$ is a subfield of $(F)_n$ by Theorem 2 in [1], since $A$ has the matrix $k$-sum $C(\tilde{f}(x))$ as its rational canonical form over $F$, where $k = n/s$. Since $m|s$, then $S_\infty(F)[A]$ has a subfield $\mathcal{M}$ of order $p^m$. Since $S_\infty(F)$ is a prime field then $\mathcal{M} = S_\infty(F)[B]$ for some $B \in S_\infty(F)[A]$. Let $h(x) \in S_\infty(F)[X]$ be the minimal polynomial of $B$ over $S_\infty(F)$. Then $h(x)$ is the minimal polynomial of $B$ over $F$, and we can choose (uniquely) $g(x) \in F[x]$ such that $\deg g(x) < s$ and $g(A) = B$.

We now sharpen the above result.

Theorem 7. Let $F = GF(q)$. Suppose $A \in (F)_n$ has minimal polynomial $f(x)$ which is prime in $F[x]$ and has degree $s$. Then for each positive divisor $m$ of $s$ and for each prime polynomial $h(x) \in F[x]$ of degree $m$, there exist precisely $m$ polynomials $g_i(x) \in F[x]$ of degrees $r_i < s$ such that $g_i(A)$ has minimal polynomial $h(x)$.

Proof. As before, the field $S_\infty(F)[A]$ has order $p^s$ and contains a subfield $S_\infty(F)[B]$ of order $p^m$. As argued in the proof of Theorem 2 in [3], any prime polynomial $h(x) \in F[x]$ of degree $m$ splits in $S_\infty(F)[B]$. The theorem follows easily.

As indicated by Section 6 of [3], there is now no difficulty in obtaining a more general result.

Theorem 8. Let $F = GF(q)$. Suppose $A \in (F)_n$ has characteristic polynomial $f(x)$ and minimal polynomial of $g(x)$ which is prime in $F[x]$. Then for each positive divisor $m$ of $n/k$ and each prime polynomial $h(x) \in F[x]$ of degree $m$, there exist precisely $m$ polynomials $g_i(x) \in F[x]$ of degrees $r_i < n/k$ such that $g_i(A)$ has characteristic polynomial $h^{m(x)}$ and minimal polynomial $h(x)$.

Our next result follows from the proof of Theorem 20 in [3].

Theorem 9. Let $F = GF(q)$, $q = p^d$. Suppose $A \in (F)_n$ is a root of a polynomial $f(x)$ which is prime in $GF[p, x]$ and has degree $s$. Then for each positive divisor $m$ of $s$, and for each polynomial $h(x)$ of degree $m$ which is prime in $GF[p, x]$ and has a root in $(F)_n$, there exist precisely $m$ polynomials $g_i(x) \in GF[p, x]$ of degrees $r_i < s$ such that $g_i(A)$ has minimal polynomial $h(x)$ over $GF(p)$.

Proof. Since $f(x)$ is prime in $GF[p, x]$ then $f(x)$ is the minimal polynomial of $A$ over $GF(p)$, and $M = S_\infty(GF(p))[A]$ is a field of order $p^s$. Hence $M$ contains a subfield $S_\infty(GF(p))[B]$ of order $p^s$ where $B \in M$. Again, any polynomial $h(x)$ of degree $m$ which is prime in $GF[p, x]$ and has a root in $(F)_n$ splits in $S_\infty(GF(p))[B]$.

We can obtain results in the "opposite direction" by modifying appropriately the constructive technique of Theorem 9 in [1] or more generally Theorem 12 in [2]. For example, consider the following

Theorem 10. Let $F = GF(q)$. Suppose $A \in (F)_n$ has characteristic polynomial $f(x)$ and minimal polynomial $f(x)$ which is prime in $F[x]$. Then for each positive integer $m$ such that $m|k$ and for each prime polynomial $h(x) \in F[x]$ of degree $mn/k$, there exist at least $mn/k$ matrices $B_i \in (F)_n$ having characteristic polynomial $h^{mn}(x)$, minimal polynomial $h(x)$, and satisfying $A = g_i(B_i)$ for unique $g_i(x) \in F[x]$ of degrees $r_i < mn/k$.

References

[3] — The number of matrix fields over GF(q), Acta Arith., this volume, pp. 315–329

UNIVERSITY OF TEXAS AT ARLINGTON
Arlington, Texas

Received on 30. 5. 1972 (201)