RANK ADDITIVITY AND MATRIX POLYNOMIALS

BY

GEORGE P.H. STYAN and AKIMICHI TAKEMURA

TECHNICAL REPORT NO. 57
SEPTEMBER 1982

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH
THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
RANK ADDITIVITY AND MATRIX POLYNOMIALS

GEORGE P.H. STYAN
McGILL UNIVERSITY

and

AKIMICHI TAKEMURA
STANFORD UNIVERSITY

TECHNICAL REPORT NO. 57
SEPTEMBER 1982

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH

Theodore W. Anderson, Project Director

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government.
Approved for public release; distribution unlimited.

Also issued as Technical Report No. 4 under U. S. Army
Research Office Contract DAAG29-82-K-0156, Department
of Statistics, Stanford University.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
RANK ADDITIVITY AND MATRIX POLYNOMIALS

Let \( A_1, \ldots, A_k \) be \( m \times n \) matrices and let \( A = \sum A_i \). Then we say that the \( A_i \)'s are rank additive to \( A \) whenever

\[
\text{rank}(A_1) + \cdots + \text{rank}(A_k) = \text{rank}(A).
\]

The earliest consideration of rank additivity may well be by Cochran (1934), who studied the distribution of quadratic forms in normal random variables. More recently, Anderson and Styan (1982), in a largely expository paper, presented various theorems on rank additivity, with particular emphasis on square matrices which are idempotent (\( A^2 = A \)), tripotent (\( A^3 = A \)) or \( r \)-potent (\( A^r = A \)). See also Khatri (1980), Takemura (1980), and Styan (1982).

In this paper we generalize some of those theorems to matrices that satisfy a general matrix polynomial equation \( P(A) = 0 \).

We begin by considering some relationships between linearly independent vector spaces, direct sums and rank additivity. There are several definitions of linear independence of vector spaces currently in use. We briefly review these and set up our notation.

**DEFINITION 1.** Let \( X \) be a (finite-dimensional) vector space and \( U_1, \ldots, U_k \) be subspaces of \( X \). \( U_1, \ldots, U_k \) are linearly independent if

\[
x_i \in U_i, \quad i = 1, \ldots, k, \quad \sum_{i=1}^{k} x_i = 0 \quad \Rightarrow \quad x_i = 0, \quad i = 1, \ldots, k.
\]

It is easy to see that \( U_1, \ldots, U_k \) are linearly independent if and only if any set of nonzero vectors \( x_i \in U_i, \quad i = 1, \ldots, k \) are linearly independent. We now list several equivalent conditions in a sequence of lemmas.

**LEMMA 1.** The vector spaces \( U_1, \ldots, U_k \) are linearly independent if and only if every vector in \( U = U_1 + \cdots + U_k \) has a unique representation in the form \( \sum_{i=1}^{k} x_i, \quad x_i \in U_i \).

**Proof:** Let \( 0 = x_1 + \cdots + x_k, \quad x_i \in U_i, \quad i = 1, \ldots, k \). Note that \( 0 \in U_i \) for all \( i \) and \( O = O + \cdots + O \). Hence by the uniqueness of the representation \( x_i = O, \quad i = 1, \ldots, k \).
Therefore $U_1, \ldots, U_k$ are independent. Conversely suppose that $U_1, \ldots, U_k$ are independent. Let $\sum_{i=1}^k x_i = \sum_{i=1}^k x_i^0$, $x_i, x_i^0 \in U_i$. Then $O = \sum_{i=1}^k x_i - x_i^0$ and $x_i - x_i^0 \in U_i$. Hence $x_i - x_i^0 = O$, $i = 1, \ldots, k$.

Rao and Yanai (1979) use the characterization in Lemma 1 as the definition of "disjointness" of the subspaces. Another definition is given by Jacobson (1953, p.28).

**LEMMA 2.** The vector spaces $U_1, \ldots, U_k$ are linearly independent if and only if

$$U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_k) = \{ O \} \quad \text{for } i = 1, \ldots, k.$$  

**Proof:** Immediate from Jacobson (1953, Th.10, p.29) and Lemma 1.

**LEMMA 3.** The vector spaces $U_1, \ldots, U_k$ are linearly independent if and only if

$$\dim(U_1 + \cdots + U_k) = \sum_{i=1}^k \dim U_i.$$  

**Proof:** Immediate from Jacobson (1953, Th.11, p.29).

If $U_1, \ldots, U_k$ are linearly independent subspaces and $U = U_1 + \cdots + U_k$ then we say that $U$ is the *direct sum* of the subspaces and denote this by

$$U = U_1 \oplus \cdots \oplus U_k = \bigoplus_{i=1}^k U_i.$$  

Consider the column space (range) $C(A_i)$ of the $m \times n_i$ matrices $A_i$, $i = 1, \ldots, k$. Let $\ell = \sum_{i=1}^k n_i$.

**LEMMA 4.** $C(A_i), i = 1, \ldots, k$ are linearly independent if and only if

$$\operatorname{rank}(A_1, A_2, \ldots, A_k) = \sum_{i=1}^k \operatorname{rank}(A_i).$$  

**Proof:** Notice that $\operatorname{rank}(A_i) = \dim C(A_i)$ and $\operatorname{rank}(A_1, \ldots, A_k) = \dim (C(A_1) + \cdots + C(A_k))$. Hence the lemma follows from Lemma 3.
Consider the $km \times m$ partitioned matrix $K_m = (I_m, \ldots, I_m)'$ and the $km \times \ell$ block diagonal matrix

$$D = \begin{pmatrix} A_1 \\ \vdots \\ A_k \end{pmatrix}.$$ 

Then Lemma 4 can be written in the form $\text{rank}(K_m'D) = \text{rank}(D)$, cf. Anderson and Styan (1982, p.8).

Now let the matrices $A_1, \ldots, A_k$ all have the same number of columns $n$. Then with $A = \sum_{i=1}^{k} A_i$ we have

**LEMMA 5.** $C(A) = \sum_{i=1}^{k} C(A_i)$ if and only if $\text{rank}(A_1, \ldots, A_k) = \text{rank}(A)$.

**Proof:** Since $C(A) \subset \sum_{i=1}^{k} C(A_i)$ always holds, $C(A) = \sum_{i=1}^{k} C(A_i)$ if and only if $\dim(C(A)) = \dim(\sum_{i=1}^{k} C(A_i))$. Now $\dim(C(A)) = \text{rank}(A)$ and $\dim(\sum C(A_i)) = \text{rank}(A_1, \ldots, A_k)$. \[\Box\]

Lemma 5 can be written in the form $\text{rank}(K_m'D) = \text{rank}(K_m'DK_n)$.

We now give the following characterization of rank additivity.

**LEMMA 6.** The matrices $A_1, \ldots, A_k$ are rank additive to $A$ if and only if $C(A) = C(A_1) \oplus \cdots \oplus C(A_k)$.

**Proof:** By Lemma 4 and Lemma 5, the column space $C(A) = C(A_1) \oplus \cdots \oplus C(A_k)$ if and only if $\text{rank}(K_m'D) = \text{rank}(D)$ and $\text{rank}(K_m'D) = \text{rank}(K_m'DK_n)$. But $\text{rank}(K_m'DK_n) \leq \text{rank}(K_m'D) \leq \text{rank}(D)$. Hence $\text{rank}(K_m'D) = \text{rank}(D)$ and $\text{rank}(K_m'D) = \text{rank}(K_m'DK_n)$ if and only if $\text{rank}(K_m'DK_n) = \text{rank}(D)$. \[\Box\]

From now on we restrict $A, A_i$ to be $n \times n$ square matrices.

**THEOREM 1.** Let $A_1, \ldots, A_k$ be square matrices, not necessarily symmetric, and let $A = \sum A_i$. Let $P(x)$ be a polynomial in the scalar $x$ with $P(0) = q$. Consider the following
statements:

(a) \( P(A_i) = 0, \ i = 1, \ldots, k, \)
(b) \( A_i A_j = 0 \) for all \( i \neq j, \)
(c) \( P(A) = 0, \)
(d) \( \sum \text{rank}(A_i) = \text{rank}(A). \)

If \( q = 0 \) then

(1) \( (b), (c), (d) \Rightarrow (a). \)

If \( q \neq 0 \) then \( P(A) = 0 \) implies that \( A \) is nonsingular and

(2) \( (b), (c), (d) \Rightarrow P(A_i) = q(I - A_i^{-1} A_i) \) and \( A_i P(A_i) = 0, \ i = 1, \ldots, k. \)

Proof: Suppose \( q = 0 \). Then (b) implies that \( O = P(A) = \sum_{i=1}^{k} P(A_i). \) Therefore for every \( x \) we obtain \( O = \sum P(A_i)x. \) Now \( P(A_i)x \in C(A_i) \). Hence by linear independence of the \( C(A_i)'s \) we have \( P(A_i)x = O \) for all \( x. \) Hence (a) holds.

Now let \( q \neq 0 \), and let the polynomial \( R(x) = xP(x). \) Then \( R(A) = 0 \) and from the previous case \( (q = 0) \) we obtain \( R(A_i) = A_i P(A_i) = 0, \ i = 1, \ldots, k. \) If \( P(A) = 0 \) then \( P(\lambda) = 0 \) for any characteristic root \( \lambda \) of \( A. \) Therefore \( q \neq 0 \) implies that 0 is not a characteristic root of \( A, \) or \( A \) is nonsingular. Then

\[
A P(A_i) = A[P(A_i) - qI] + qA \\
= A_i[P(A_i) - qI] + qA \\
= q(A - A_i),
\]

from which (2) follows at once.

When the polynomial

\( P(x) = P_2(x) = x^2 - x, \)

then (1) may be strengthened to

\( (c)_{2}, (d) \leftrightarrow (a), (b), \)

- 4 -
where \((c)_2\) is \((c)\) with \(P = P_2\). This is Cochran's Theorem (cf. Anderson and Styan, 1982, Th.1.1). When

\[ P(x) = P_3(x) = x^3 - x, \]

then (1) may be strengthened

\[(c)_3, (d), (e) \Leftrightarrow (a), (b), \]

where

\[(e) \quad \mathbf{A} \mathbf{A}_i = \mathbf{A}_i \mathbf{A}, \quad i = 1, \ldots, k, \]

cf. Anderson and Styan (1982, Th.3.1). Here \((c)_3\) is \((c)\) with \(P = P_3\). Takemura (1980, Th.3.2) showed that (3) still holds with \((c)_3\) replaced by \((c)_r\) for \(P(x) = P_r(x) = x^r - x\).

Notice that the polynomials \(P_2, P_3\) and \(P_r\) have no multiple root; we obtain further results when the polynomial \(P\) has no multiple root. First we show that there exists a “nullity-additivity” relation underlying a matrix polynomial with no multiple root. Anderson and Styan (1982, p.5) showed that

\[(4) \quad \nu(A - A^2) = \nu[A(I - A)] = \nu(A) + \nu(I - A), \]

where

\[(5) \quad \nu(A) = n - \text{rank}(A) \]

is the (column) nullity of the \(n \times n\) matrix \(A\).

Equation (4) is a special case of equality in Sylvester's law of nullity:

\[ \nu(AB) \leq \nu(A) + \nu(B), \]

where \(A\) is \(m \times n\) and \(B\) is \(n \times \ell\), say. Then

\[(6) \quad \nu(AB) = \nu(A) + \nu(B) \]

if and only if

\[(7) \quad \mathcal{N}(A) \subseteq \mathcal{C}(B), \]

where \(\mathcal{N}(\cdot)\) denotes null space, cf. Satake (1975, p.124). See also Marsaglia and Styan (1974, p.275). Using this fact we obtain
THEOREM 2. Let $A$ be a square matrix and let $x_1, \ldots, x_d$ be distinct scalars. Then

$$
\nu[\prod_{i=1}^d (A - x_i I)] = \sum_{i=1}^d \nu(A - x_i I).
$$

Proof: Let $u \in N(A - x_1 I)$. Then $Au = x_1 u$ and since $\prod_{i=2}^d (x_1 - x_i) \neq 0$ we see that

$$
u = \prod_{i=2}^d (A - x_i I)u / \prod_{i=2}^d (x_1 - x_i) \in C[\prod_{i=2}^d (A - x_i I)]
$$

and so

$$
\nu[\prod_{i=1}^d (A - x_i I)] = \nu(A - x_1 I) + \nu[\prod_{i=2}^d (A - x_i I)],
$$

since (6)$\Leftrightarrow$(7). Repeating this argument $d - 2$ times establishes (8).

Theorem 2 yields the following corollaries:

COROLLARY 1. Let the polynomial $P$ have degree $d$ and distinct roots $x_1, \ldots, x_d$, and let the matrix $A$ be $n \times n$. Then

$$
\nu[P(A)] = \nu[\prod_{i=1}^d (A - x_i I)] = \sum_{i=1}^d \nu(A - x_i I).
$$

Moreover,

$$
P(A) = 0 \Leftrightarrow \sum_{i=1}^d \nu(A - x_i I) = n
$$

$$
\Leftrightarrow \sum_{i=1}^d \text{rank}(A - x_i I) = (d - 1)n,
$$

and the set $\{x_1, \ldots, x_d\}$ contains all distinct characteristic roots of $A$.

Proof: Equation (9) follows from $P(A) = 0 \Leftrightarrow \nu[P(A)] = n$ and from (5). If $P(A) = 0$ then any characteristic root of $A$ is a root of $P$. Hence $\{x_1, \ldots, x_d\}$ contains all distinct characteristic roots of $A$. 

- 6 -
COROLLARY 2. Let \( \omega = \exp[2\pi i/(r-1)] \), where the integer \( r \geq 2 \) and let the matrix \( A \) be \( n \times n \). Then

\[
\nu(A - A^r) = \nu(A) + \nu(I - A) + \sum_{s=1}^{r-2} \nu(\omega^s I - A),
\]

and

\[
A = A^r \iff \nu(A) + \nu(I - A) + \sum_{s=1}^{r-2} \nu(\omega^s I - A) = n
\]

\[
\iff \text{rank}(A) + \text{rank}(I - A) + \sum_{s=1}^{r-2} \text{rank}(\omega^s I - A) = (r - 1)n.
\]

When \( r = 2 \) the summation in Corollary 2 disappear and (10) reduces to (4).

When \( r = 3 \), equation (10) becomes

\[
\nu(A - A^3) = \nu(A) + \nu(I - A) + \nu(I + A),
\]


Another consequence of \( P \) having no multiple root is the diagonability of the matrix \( A \) which satisfies \( P(A) = 0 \).

LEMMA 7. The square matrix \( A \) is diagonable if and only if there exists a polynomial \( P \) with no multiple root such that \( P(A) = 0 \).

A matrix \( A \) is said to be diagonable if there exists a nonsingular \( F \) such that \( F^{-1}AF \) is diagonal, and then the minimal polynomial has no multiple root (cf. e.g., Mirsky, 1955, Th.10.2.5, p.297). The polynomial \( P \) in Lemma 7 must be a multiple of (or actually) the minimal polynomial. Lemma 7 shows that an idempotent, tripotent or \( r \)-potent matrix \( A \) is diagonable.

We may prove Lemma 7 using the algebraic and geometric multiplicities of the (distinct) characteristic roots \( \lambda_1, \ldots, \lambda_p \) of \( A \). Let \( am_j, j = 1, \ldots, p \) denote the algebraic multiplicity of \( \lambda_j \), namely the multiplicity of \( \lambda_j \) as a root of the characteristic equation. Let \( gm_j, j = 1, \ldots, p \), denote the geometric multiplicity of \( \lambda_j \), namely the nullity \( \nu(A - \lambda_j I) \). Note that \( am_j \geq gm_j, j = 1, \ldots, p \). (See e.g. Mirsky 1955, p.294). The characteristic root \( \lambda_j \) is said to be regular if \( am_j = gm_j \).
LEMMA 8. The square matrix $A$ is diagonalizable if and only if all its characteristic roots are regular.

Proof: See e.g. Mirsky (1955, Th.10.2.3). 

Proof of Lemma 7. Let $P(x) = (x - x_1)(x - x_2) \cdots (x - x_d)$, where $d = \deg P$, and $x_1, \ldots, x_d$ are the distinct roots of $P(x) = 0$ and suppose $P(A) = 0$. Then

$$O = P(A) = (A - x_1 I)(A - x_2 I) \cdots (A - x_d I).$$

Define $g m_i^* = gm_j$ if $x_i = \lambda_j$ for some $j$ and $g m_i^* = 0$ otherwise. Then $\nu(A - x_i I) = g m_i^*$ for all $i$. [Note that is $x_i$ is not a characteristic root of $A$ then $A - x_i I$ is nonsingular and $\nu(A - x_i I) = 0 \Rightarrow g m_i^*$.] Then by Theorem 2

$$n = \sum_{i=1}^{d} \nu(A - x_i I) = \sum_{i=1}^{d} g m_i^* \leq \sum_{j=1}^{p} g m_j \leq \sum_{j=1}^{p} a m_j = n.$$

Hence the inequalities above collapse and we have $a m_i = g m_i$, $i = 1, \ldots, p$. By Lemma 8 $A$ is diagonalizable.

To go the other way let $A$ be diagonalizable. Then we may write

$$A = F A F^{-1} = F \text{diag}(\lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_p) F^{-1},$$

where $\lambda_1, \ldots, \lambda_p$ are distinct roots of $A$. Let $P(x) = (x - \lambda_1) \cdots (x - \lambda_p)$. Then

$$P(A) = F P(A) F^{-1} = F (A - \lambda_1 I) \cdots (A - \lambda_p I) F^{-1} = F O F^{-1} = 0,$$

and the result is established. 

The two matrices $A, B$ are said to be simultaneously diagonalizable if $A, B$ can be diagonalized by the same nonsingular matrix $F$. We then have the following extension of Theorem 1.
THEOREM 3. Let $A_1, \ldots, A_k$ be $n \times n$ matrices, not necessarily symmetric, and let $A = \sum A_i$ be diagonalizable. Suppose that

(b) $A_i A_j = 0$ for all $i \neq j$,

(d) $\sum \text{rank}(A_i) = \text{rank}(A)$.

Then $A, A_1, \ldots, A_k$ are all simultaneously diagonalizable and for some nonsingular $F$

(11) $F^{-1} A F = \text{diag}(\lambda_{j(1)}, \ldots, \lambda_{j(n)})$,

and

(12) $F^{-1} A_i F = \text{diag}(0, \ldots, 0, \lambda_{j(1)}, \ldots, \lambda_{j(n)}, 0, \ldots, 0)$,

where $r_i = \text{rank}(A_i)$, $j(i) \in \{1, \ldots, p\}$, $i = 1, \ldots, n$.

Proof: Let $P(x)$ be a polynomial with no multiple root and such that $P(A) = 0$. If $P(0) = 0$ then by Theorem 1 $P(A_i) = 0$ and hence $A_i$ is diagonalizable, $i = 1, \ldots, k$. If $P(0) \neq 0$ then 0 is not a root of $P(x) = 0$. Hence $R(x) = xP(x)$ still has no multiple root. By Theorem 1 again $R(A_i) = 0$ and hence $A_i$ is diagonalizable, $i = 1, \ldots, k$. In any event $A_1, \ldots, A_k$ are diagonalizable. From (b) it follows that $A, A_1, \ldots, A_k$ are simultaneously diagonalizable (by $F$). See e.g., Mirsky (1955, Th.10.6.3., p.318) or Takemura (1980, Th.4.3).

Let $F^{-1} A F = A_1$, $F^{-1} A_i F = A_i$. Then $A_i A_j = 0$ for all $i \neq j$, $A = \sum A_i$, $\text{rank}(A) = \sum \text{rank}(A_i)$ imply that by rearranging the coordinates (if necessary) we can obtain (11) and (12). \qed

We extend Theorem 3 with:

THEOREM 4. Let $A_1, \ldots, A_k$ be $n \times n$ matrices, not necessarily symmetric, and let $A = \sum A_i$. Suppose that

(b) $A_i A_j = 0$ for all $i \neq j$.

Then the set of nonzero characteristic roots of $A$ coincides with the set of all the nonzero characteristic roots of all the $A_i$, $i = 1, \ldots, k$. Furthermore the nonzero characteristic
root $\lambda$ of $A$ is regular if and only if $\lambda$ is a regular nonzero characteristic root of each $A_i$, $i = 1, \ldots, k$.

If in addition

$$(d) \quad \sum_{i=1}^{k} \text{rank}(A_i) = \text{rank}(A),$$

then the characteristic root 0 of $A$ is regular if and only if 0 is a regular characteristic root of each $A_i$, $i = 1, \ldots, k$. Equivalently

$$\text{rank}(A^2) = \text{rank}(A) \iff \text{rank}(A_i^2) = \text{rank}(A_i), \quad i = 1, \ldots, k.$$ 

Mäkeläinen and Styan (1976, Lemma 2) have shown that the zero characteristic root of a matrix $A$ is regular if and only if $\text{rank}(A^2) = \text{rank}(A)$. Such a matrix $A$ is said to have index 1, cf. Ben-Israel and Greville (1974, p.169). Marsaglia and Styan (1974, Th.15, p.286) proved that if $A_1, \ldots, A_k$ all have index 1 then (b)$\Rightarrow$(d). From Lemma 7 and 8 it follows that when the polynomial $P$ in Theorem 1 has no multiple root and $P(0) = 0$ then (a), (b) $\Rightarrow$ (c), (d).

Proof of Theorem 4. Let $A$ have rank $r$, and let $A_i$ have rank $r_i$, $i = 1, \ldots, k$. Let $\lambda_1, \ldots, \lambda_\ell$ be the nonzero characteristic roots of $A$. Let $m_{ij}$ be the algebraic multiplicity and $g_{ij}$ the geometric multiplicity of $\lambda_j$ as a characteristic root of $A_i$, so that, cf. e.g., Mirsky (1955, Th.7.6.1, p.214),

$$n \geq m_{ij} \geq g_{ij} \geq 0; \quad i = 1, \ldots, k, \quad j = 1, \ldots, \ell.$$  

Then $\lambda_j$ is a regular characteristic root of $A_i$ whenever $m_{ij} = g_{ij}$. [Notice that $g_{ij} = 0 \iff m_{ij} = 0$; we will then speak of $\lambda_j$ as a regular characteristic root of $A_i$ even though $A_i$ does not have $\lambda_j$ as a root.] Let $m_{0j}$ be the algebraic multiplicity and $g_{0j}$ the geometric multiplicity of $\lambda_j$ as a characteristic root of $A$. Let $m_{i0}$ be the algebraic multiplicity and $g_{i0}$ the geometric multiplicity of 0 as a characteristic root of $A_i$, $i = 1, \ldots, k$. Then

$$m_{i0} = n - m_i = n - \sum_{j=1}^{\ell} m_{ij}, \quad g_{i0} = n - r_i = n - \text{rank}(A_i), \quad i = 1, \ldots, k.$$ 

- 10 -
Hence \( n \geq r_i \geq m_i \geq 0; \quad i = 1, \ldots, k. \) Let \( m_{00} \) be the algebraic multiplicity and let \( g_{00} \) be the geometric multiplicity of 0 as a characteristic root of \( A. \) Then

\[
\begin{align*}
m_{00} &= n - m_0 = n - \sum_{j=1}^{\ell} m_{0j}, \\
g_{00} &= n - r = n - \text{rank}(A).
\end{align*}
\]

Hence \( n \geq r \geq m_0 \geq 0. \)

Let \( A_i = B_i C_i', \quad i = 1, \ldots, k, \) be full rank decompositions, with \( B_i \) and \( C_i \) both \( n \times r_i \) of rank \( r_i. \) Then \( A = \sum_1^k A_i = \sum_1^k B_i C_i' = BC', \) where

\[
B = (B_1, \ldots, B_k) \quad \text{and} \quad C = (C_1, \ldots, C_k)
\]

are both \( n \times \sum_1^k r_i. \)

Now suppose that (b) holds. Then \( C_i' B_j = 0 \) for all \( i \neq j \) and so \( C'B = \text{diag}(C_1' B_1, \ldots, C_k' B_k) \) is a block diagonal matrix.

Let \( am_j(A) \) denote the algebraic multiplicity and \( gm_j(A) \) denote the geometric multiplicity of \( \lambda_j \) as a characteristic root of \( A. \) Then since the matrices \( FG \) and \( GF \) have the same nonzero characteristic roots (cf. e.g., Mirsky, 1955, Th.7.2.3., p.200), we may write for \( j = 1, \ldots, \ell \)

\[
m_{0j} = am_j(A) = am_j(BC') = am_j(C'B) = \sum_{i=1}^{k} am_j(C_i' B_i) = \sum_{i=1}^{k} am_j(B_i C_i') = \sum_{i=1}^{k} m_{ij},
\]

while

\[
m_0 = \sum_{j=1}^{\ell} m_{0j} = \sum_{i=1}^{k} \sum_{j=1}^{\ell} m_{ij} = \sum_{i=1}^{k} m_i.
\]

so that \( n - am_0(A) = \sum_1^k [n - am_0(A_i)]. \)

We now use the result that the matrices \( FG - I \) and \( GF - I \) have the same nullity, cf. Ouellette (1981, p.246). Then for each \( j = 1, \ldots, \ell \)

\[\]
\[ g_{ij} = gm_j(A) = \nu(A - \lambda_j I) = \nu(BC' - \lambda_j I) = \nu(C'B - \lambda_j I) \]
\[ = \sum_{i=1}^{k} \nu(C'_i B_i - \lambda_j I_{r_i}) = \sum_{i=1}^{k} \nu(B_i C'_i - \lambda_j I_n) \]
\[ = \sum_{i=1}^{k} \nu(A_i - \lambda_j I_n) = \sum_{i=1}^{k} g_{ij}. \]

Hence when (b) holds all the nonzero characteristic roots of the \( A_i, \ i = 1, \ldots, k, \) must be characteristic roots of \( A, \) and all the nonzero characteristic roots of \( A \) must be characteristic roots of \( A_i \) for some \( i. \)

Furthermore, since \( g_{ij} \leq m_{ij} \) from (13) we obtain
\[ g_{0j} = \sum_{i=1}^{k} g_{ij} \leq \sum_{i=1}^{k} m_{ij} = m_{0j}; \quad j = 1, \ldots, \ell, \]
and so for each \( j = 1, \ldots, \ell \)
\[ g_{0j} = m_{0j} \Leftrightarrow g_{ij} = m_{ij}; \quad i = 1, \ldots, k. \]

Thus when (b) holds, \( \lambda_j \) is a regular nonzero characteristic root of \( A \) if and only if \( \lambda_j \) is also a regular characteristic root of each \( A_i, \quad i = 1, \ldots, k. \)

Now suppose that both (b) and (d) hold. Then substitution in
\[ n - gm_0(A) = \tau \leq \sum_{i=1}^{k} r_i = \sum_{i=1}^{k} [n - gm_0(A_i)] \]
yields \( n - gm_0(A) = \sum_{i=1}^{k} [n - gm_0(A_i)] \) and so
\[ gm_0(A) = n - \sum_{i=1}^{k} [n - gm_0(A_i)] \leq n - \sum_{i=1}^{k} [n - am_0(A_i)] = n - \sum_{i=1}^{k} m_i = n - m_0. \]

Hence 0 is a regular characteristic root of \( A \) if and only if 0 is a regular characteristic root of \( A_i, \) for all \( i = 1, \ldots, k. \)
ACKNOWLEDGEMENTS

Part of this research was completed while the first author was visiting the Department of Statistics at Stanford University in June 1981 and June 1982. He is most grateful to Professor T.W. Anderson for his hospitality. This research was supported in part by Natural Sciences and Engineering Research Council of Canada, Grant No. A7274, Gouvernement du Québec, Programme de chercheurs et d’action concertée, subvention no. EQ-961, Office of Naval Research Contract N00014-75-C-0442, and Army Research Office Contract DAAG29-82-K-0156.

REFERENCES


Marsaglia, George, and Styan, George P.H. (1974). Equalities and inequalities for ranks of


17. "General Exponential Models for Discrete Observations,"

18. "On the Interrelationships among Sufficiency, Total Sufficiency and


Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models

    Raul Pedro Mentz, September 8, 1975.

22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting,"
    Mituaki Huzii, February 1976.

23. "Estimating Means when Some Observations are Classified by Linear
    Discriminant Function," Chien-Pai Han, April 1976.


    September 1976.


28. "Estimation of the Parameters of Finite Location and Scale Mixtures,"
    Javad Behboodian, October 1976.

29. "Identification of Parameters by the Distribution of a Maximum

30. "Discrimination Between Stationary Gaussian Processes, Large Sample

31. "Principal Components in the Nonnormal Case: The Test for Sphericity,"
    Christine M. Waternaux, October 1977.

32. "Nonnegative Definiteness of the Estimated Dispersion Matrix in a
TECHNICAL REPORTS (continued)


57. "Rank Additivity and Matrix Polynomials," George P.H. Styan and Akimichi Takemura September 1982
### Report Documentation Page

<table>
<thead>
<tr>
<th>1. Report Number</th>
<th>2. GOVT Accession No.</th>
<th>3. Recipient's Catalog Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>57</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Title (and Subtitle)</th>
<th>5. Type of Report &amp; Period Covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank Additivity and Matrix Polynomials</td>
<td>Technical Report</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7. Author(s)</th>
<th>8. Contract or Grant Number(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>George P.H. Styan Akimichi Takemura</td>
<td>N00014-75-C-0442</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9. Performing Organization Name and Address</th>
<th>10. Program Element, Project, Task Area &amp; Work Unit Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department of Statistics - Sequoia Hall</td>
<td>(NR-042-034)</td>
</tr>
<tr>
<td>Stanford University</td>
<td></td>
</tr>
<tr>
<td>Stanford, California 94305</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11. Controlling Office Name and Address</th>
<th>12. Report Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Office of Naval Research</td>
<td>September 1982</td>
</tr>
<tr>
<td>Statistics and Probability Program Code 436</td>
<td></td>
</tr>
<tr>
<td>Arlington, Virginia 22217</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>14. Monitoring Agency Name &amp; Address (if different from Controlling Office)</th>
<th>15. Security Class. (of this report)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UNCLASSIFIED</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>16. Distribution Statement (of this Report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approved for public release; distribution unlimited.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>17. Distribution Statement (of the abstract entered in Block 20, if different from Report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Also issued as Technical Report No. 4 under U. S. Army Research Contract No. DAAG29-82-K-0156, Department of Statistics, Stanford University.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>18. Supplementary Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>19. Key Words (Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cochran's theorem, rank additivity, matrix polynomials, diagonability, regular characteristic root.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>20. Abstract (Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Various results are given concerning the logical relation between the rank additivity condition of matrices and polynomial equations satisfied by these matrices, generalizing earlier results on idempotent, tripotent, and r-potent matrices.</td>
</tr>
</tbody>
</table>