SPACES OF MATRICES WITH SEVERAL ZERO EIGENVALUES

M. D. ATKINSON

Let $V$ be an $n$-dimensional vector space over some field $F$, $|F| \geq n$, and let $\mathcal{X}$ be a space of linear mappings from $V$ into itself ($\mathcal{X} \subseteq \text{Hom}(V, V)$) with the property that every mapping has at least $r$ zero eigenvalues. If $r = 0$ this condition is vacuous but if $r = 1$ it states that $\mathcal{X}$ is a space of singular mappings; in this case Flanders [1] has shown that $\dim \mathcal{X} \leq n(n - 1)$ and that if $\dim \mathcal{X} = n(n - 1)$ then either $\mathcal{X} = \text{Hom}(V, U)$ for some $(n - 1)$-dimensional subspace $U$ of $V$ or there exists $v \in V, v \neq 0$ such that $\mathcal{X} = \{X \in \text{Hom}(V, V): vX = 0\}$. At the other extreme $r = n$, the case of nilpotent mappings, a theorem of Gerstenhaber [2] states that $\dim \mathcal{X} \leq \frac{1}{2}n(n - 1)$ and that if $\dim \mathcal{X} = \frac{1}{2}n(n - 1)$ then $\mathcal{X}$ is the full algebra of strictly lower triangular matrices (with respect to some suitable basis of $V$).

In this paper we intend to derive similar results for a general value of $r$ thereby providing some common ground for the above theorems. The condition on the cardinal of $F$ plays an important role in our proofs although we do not know whether it is essential to the theorems. Flanders required also the constraint $\text{char } F \neq 2$ at one point in his proof; we do not require this and in fact our type of argument allows the field in Flanders' theorem to have arbitrary characteristic.

Theorem 1. Let $V$ be an $n$-dimensional vector space over some field $F$, $|F| \geq n$, and let $\mathcal{X}$ be a subspace of $\text{Hom}(V, V)$. Suppose that every mapping in $\mathcal{X}$ has at least $r$ zero eigenvalues. Then $\dim \mathcal{X} \leq \frac{1}{2}(r - 1) + n(n - r)$.

Proof. Gerstenhaber's theorem proves our result in the case $n - r = 0$ and provides the first step of an induction on $n - r$. Assume now that $n > r > 0$ and that the theorem holds for spaces whose transformations all have more than $r$ zero eigenvalues. Since $\frac{1}{2}(r + 1) + n(n - r - 1) < \frac{1}{2}(r - 1) + n(n - r)$ we can suppose that $\mathcal{X}$ contains a transformation with precisely $r$ zero eigenvalues. By an appropriate choice of basis for $V$ we may take this to be (represented by)

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

with $A_1$ an $r \times r$ nilpotent matrix and $A_2$ non-singular. By taking $A_1$ in rational canonical form we may take it to be strictly lower triangular. From now on we shall take $\mathcal{X}$ to consist of $n \times n$ matrices $\begin{pmatrix} J & K \\ M & N \end{pmatrix}$ partitioned so that $J$ is an $r \times r$ matrix.

Consider any matrix $B \in \mathcal{X}$ of the form $\begin{pmatrix} J & K \\ 0 & N \end{pmatrix}$. For all $\alpha \in F$ $\begin{pmatrix} J + \alpha A_1 & K \\ 0 & N + \alpha A_2 \end{pmatrix}$ has at least $r$ zero eigenvalues and, except for at most $n - r$ eigenvalues of $-A_2^{-1}N$, $N + \alpha A_2$ is non-singular. Hence for all but at most $n - r$ values of $\alpha$, $J + \alpha A_1$ is

Received 24 May, 1979.

[BULL. LONDON MATH. SOC., 12 (1980), 89–95]
nilpotent. However the coefficients of the characteristic polynomial of \( J + \alpha A_1 \) are polynomials in \( \alpha \) which, since \( A_1 \) is singular, are of degree at most \( r - 1 \). The non-leading coefficients of the characteristic polynomial all vanish for at least \( r \) values of \( \alpha \) (since \( |F| \geq n \)) and so are identically zero i.e. \( J + \alpha A_1 \) is nilpotent for all \( \alpha \) and in particular \( J \) is nilpotent. Let \( \mathcal{J} \) be the space of nilpotent \( r \times r \) matrices which arise in this way; then \( \dim \mathcal{J} \leq \frac{1}{2} r(r-1) \).

Let \( J_1, \ldots, J_r \) be a basis for \( \mathcal{J} \) and consider the \( r^2 \) \( t \)-vectors \( l_{pq} \) whose \( k \)-th component is \( (J_k)_{pq} \). Since \( t \leq \frac{1}{2} r(r-1) \) we may choose a set \( L \) of \( \frac{1}{2} r(r-1) \) index pairs \((p, q)\) such that \( \{l_{pq} : (p, q) \in L\} \) spans the space generated by all the \( l_{pq} \). Let \( U \) be the complementary set of \( r \times r \) matrix \( J \) may be written uniquely as \( J = J_L + J_U \) where \( (J_L)_{pq} = 0 \) if \((p, q) \in U\) and \( (J_U)_{pq} = 0 \) if \((p, q) \in L\). By the choice of \( L \) we have the condition \( J \in \mathcal{J} \) and \( J_L = 0 \) implies \( J = 0 \), or equivalently

\[
\begin{pmatrix}
J_L + J_U & K \\
0 & N
\end{pmatrix} \in \mathcal{X} \quad \text{and} \quad J_L = 0 \quad \text{implies} \quad J_U = 0. \quad (*)
\]

By expressing each matrix of \( \mathcal{X} \) in the form \( \begin{pmatrix} J_L + J_U & K \\ M & N \end{pmatrix} \) we may define a linear map \( \phi \) on \( \mathcal{X} \) by

\[
\begin{pmatrix}
J_L + J_U & K \\
M & N
\end{pmatrix} \xrightarrow{\phi} \begin{pmatrix} J_L & 0 \\ M & N \end{pmatrix}.
\]

Then, by \((*)\), \( \left\{ \begin{pmatrix} 0 \\ K \end{pmatrix} \right\} \in \mathcal{X} \) is the kernel of \( \phi \). Let

\[
\mathcal{K} = \left\{ K : \begin{pmatrix} 0 & K \\ 0 & 0 \end{pmatrix} \in \mathcal{X} \right\}, \quad \mathcal{L} = \left\{ J_L : \begin{pmatrix} J & K \\ M & N \end{pmatrix} \in \mathcal{X} \right\},
\]

\[
\mathcal{M} = \left\{ M : \begin{pmatrix} J & K \\ M & N \end{pmatrix} \in \mathcal{X} \right\}, \quad \mathcal{N} = \left\{ N : \begin{pmatrix} J & K \\ M & N \end{pmatrix} \in \mathcal{X} \right\}.
\]

Clearly \( \dim \mathcal{X} = \dim \text{range}(\phi) + \dim \text{kernel}(\phi) \)

\[
\leq \dim \mathcal{L} + \dim \mathcal{M} + \dim \mathcal{N} + \dim \mathcal{K}
\]

\[
\leq \frac{1}{2} r(r-1) + (n-r)^2 + \dim \mathcal{M} + \dim \mathcal{K}
\]

and the proof will be completed when we have shown that \( \dim \mathcal{M} + \dim \mathcal{K} \leq r(n-r) \).

To do this we shall construct a non-singular bilinear form \( \theta : \mathcal{A}_{r,n-r} \times \mathcal{A}_{r,n-r} \rightarrow F \) where \( \mathcal{A}_{r,n-r} \) denote the full spaces of \( r \times (n-r) \), \( (n-r) \times r \) matrices over \( F \) and show that \( \theta \) vanishes on \( \mathcal{X} \times \mathcal{M} \). This will prove that \( \mathcal{M} \) is contained in the annihilator of \( \mathcal{K} \) which has dimension \( r(n-r)-\dim \mathcal{K} \), giving \( \dim \mathcal{M} \leq r(n-r)-\dim \mathcal{K} \) as required.

Consider any matrices \( K \in \mathcal{K} \), \( M \in \mathcal{M} \). Then there exist \( Y = \begin{pmatrix} 0 & K \\ 0 & 0 \end{pmatrix} = (y_{ij}) \in \mathcal{X} \)

and \( X = \begin{pmatrix} J & H \\ M & N \end{pmatrix} = (x_{ij}) \in \mathcal{X} \). Since \( X + \alpha A + \beta Y \) has at least \( r \) zero eigenvalues for all
\( \alpha, \beta \) the coefficient of \( \alpha^{n-r-1}\beta^{r-1} \) in \( \det (X + \alpha A + \beta Y - \lambda I) \) is zero (see [2] p. 615). The total coefficient of \( \lambda^{r-1} \) in this determinant is (apart from a sign) the sum of all \((n-r+1)\)-rowed principal minors of \( X + \alpha A + \beta Y \) and we require the coefficient of \( \alpha^{n-r-1}\beta \) in this sum. Consider a typical one of these minors

\[
M(S, T) = \det \begin{pmatrix} J' + \alpha A'_1 & H' + \beta K' \\ M' & N' + \alpha A'_2 \end{pmatrix}
\]

where \( J' + \alpha A'_1 \) is the principal submatrix of \( X + \alpha A + \beta Y \) whose rows are indexed by the set \( S \subseteq \{1, 2, \ldots, r\} \) and \( N' + \alpha A'_2 \) is the principal submatrix whose rows are indexed by the set \( T \subseteq \{ r+1, \ldots, n \} \) and where \( |S \cup T| = n-r+1 \). By direct calculation the coefficient of \( \alpha^{n-r-1}\beta \) in this minor is, to within a sign,

\[
\sum_{i,j \in S, s \in T} C_{ij} D_{st} y_{is} x_{sj}
\]

where \( (C_{ij}) \) is the matrix of cofactors of \( A'_1 \) and \( (D_{st}) \) is the matrix of cofactors of \( A'_2 \). We shall express such a coefficient as

\[
y Z x^t
\]

where

\[
y = (y_{1, r+1}, \ldots, y_{1n}, y_{2, r+1}, \ldots, y_{2n}, \ldots, y_{r, r+1}, \ldots, y_{rn}),
\]

\[
x = (x_{r+1, 1}, \ldots, x_{r+1, n}, x_{r+2, 1}, \ldots, x_{r+2, n}, \ldots, x_{r+1, r}, \ldots, x_{n}),
\]

and \( Z \) is an \( r(n-r) \times r(n-r) \) matrix whose entries depend upon \( (C_{ij}) \) and \( (D_{st}) \).

If \( |S| > 1 \) then, since \( A'_1 \) is strictly lower triangular, \( C_{ij} = 0 \) for \( i, j \in S, i \geq j \) and the matrix \( Z \) then has the form

\[
\begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix}
\]

with zeros above the diagonal and where each diagonal 0 represents an \( (n-r) \times (n-r) \) matrix of zeros.

If \( |S| = 1 \), say \( S = \{i\} \), then \( |T| = n-r \), \( C_{ii} = 1 \) and \( (D_{si}) = E \) is the matrix of cofactors of \( A_2 \). In this case the matrix \( Z \) has the form

\[
\begin{pmatrix}
0 \\
0 \\
\vdots \\
E \cdot \\
0
\end{pmatrix}
\]
where each diagonal block is an \( (n - r) \times (n - r) \) matrix, \( E \) is the \( i \)-th diagonal block and there are zeros elsewhere.

Summing over all appropriate \( S, T \) we obtain the coefficient of \( x^{n-r-1} \beta^t \) in the form \( yZx^t \) where

\[
Z = \begin{pmatrix}
E & & \\
& E & \\
& & \ddots \\
& & & E
\end{pmatrix}
\]

and has zeros above the diagonal. Since \( A_2 \) is non-singular so also is \( E \) and thus \( Z \) is non-singular. This matrix \( Z \) defines our non-singular bilinear form vanishing on \( \mathcal{X} \times \mathcal{M} \) and so completes the proof of the theorem.

In the next theorem we consider the case where \( \dim \mathcal{X} \) is maximal and determine the possibilities for \( \mathcal{X} \). We first prove

**Lemma.** Let \( V \) be an \( n \)-dimensional vector space over some field \( F \), \( |F| \geq n \), and let \( \mathcal{X} \) be a subspace of \( \text{Hom}(V, V) \). Suppose that \( W_1 \) and \( W_2 \) are \( \mathcal{X} \)-invariant subspaces of \( V \) such that

(i) \( \dim W_1 = s_1 \), \( \dim V/W_2 = s_2 \), \( \dim W_1 \cap W_2 = t \),

(ii) \( \mathcal{X} \) induces spaces of nilpotent mappings in each of \( W_1 \) and \( V/W_2 \),

(iii) every transformation in \( \mathcal{X} \) has at least \( r \) zero eigenvalues.

Then \( \dim \mathcal{X} \leq \frac{1}{2}r(r-1) + n(n-r) - (s_1-t)(n-s_2-t) \).

**Proof.** \( W_1 + W_2 \) has dimension \( \dim W_1 + \dim W_2 - \dim W_1 \cap W_2 = s_1 + n - s_2 - t \) and we may choose a basis of it having the form

\[
e_1, \ldots, e_t, f_1, \ldots, f_{n-t}, g_1, \ldots, g_{n-s_2-t}.
\]

such that \( e_1, \ldots, e_t \) is a basis for \( W_1 \cap W_2 \), \( e_t, f_1, \ldots, f_{n-t} \) is a basis for \( W_1 \), and \( e_1, \ldots, e_t, g_1, \ldots, g_{n-s_2-t} \) is a basis for \( W_2 \). We complete this to a basis of \( V \) with \( h_1, \ldots, h_{2n-t+1} \).

With respect to this basis each transformation of \( \mathcal{X} \) has matrix

\[
\begin{pmatrix}
N_1 & 0 & 0 & 0 \\
A_1 & N_2 & 0 & 0 \\
A_2 & 0 & B & 0 \\
A_3 & A_4 & A_5 & N_3
\end{pmatrix}
\]

\( N_1 \) and \( N_2 \) are nilpotent \( t \times t \) and \( (s_1-t) \times (s_1-t) \) matrices since \( \mathcal{X} \) restricted to \( W_1 \) is a space of nilpotent mappings. \( N_3 \) is a nilpotent \( (s_2 - s_1 + t) \times (s_2 - s_1 + t) \) matrix since \( \mathcal{X} \) induces nilpotent transformations on \( V/(W_1 + W_2) \). Since the whole matrix has at least \( r \) zero eigenvalues the \( (n-s_2-t) \times (n-s_2-t) \) matrix \( B \) has at least \( r - s_2 - t \) zero
eigenvalues. Now, by Theorem 1, the space of all \( \begin{pmatrix} N_1 & 0 \\ A_1 & N_2 \end{pmatrix} \) has dimension at most \( \frac{1}{2}s_1(s_1-1) \), the space of all matrices \( N_3 \) has dimension at most
\[
\frac{1}{2}(s_2-s_1+t)(s_2-s_1+t-1),
\]
and the space of all matrices \( B \) has dimension at most
\[
\frac{1}{2}(r-s_2-t)(r-s_2-t-1)+(n-s_2-t)(n-r).
\]
A routine calculation now shows that
\[
\dim \mathcal{X} \leq \frac{1}{2}r(r-1) + n(n-r) - (s_1-t)(n-s_2-t).
\]

**Theorem 2.** Let \( V \) be an \( n \)-dimensional vector space over some field \( F, |F| \geq n \), and let \( \mathcal{X} \) be a subspace of \( \text{Hom}(V, V) \). Suppose that every mapping in \( \mathcal{X} \) has at least \( r \) zero eigenvalues and that \( \mathcal{X} \) has dimension precisely \( \frac{1}{2}r(r-1) + n(n-r) \). Then \( V \) contains \( \mathcal{X} \)-invariant subspaces \( V_1, V_2 \) with \( V_1 \leq V_2 \) and \( \dim V_1 + \dim V/V_2 = r \) and such that \( \mathcal{X} \) operates on each of \( V_1 \) and \( V/V_2 \) as the full algebra of strictly lower triangular matrices (with respect to suitable bases). Moreover the subspaces \( V_1, V_2 \) are uniquely determined by \( \mathcal{X} \) unless \( r = n \).

**Proof.** We continue to use the machinery and notation \( A, L, U, \phi, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{K}, \theta \) introduced in the proof of Theorem 1.

From
\[
\frac{1}{2}r(r-1) + n(n-r) = \dim \mathcal{X} \leq \dim \mathcal{L} + \dim \mathcal{N} + \dim \mathcal{M} + \dim \mathcal{K}
\]
we obtain
\[
\dim \mathcal{L} = \frac{1}{2}r(r-1), \quad \dim \mathcal{N} = (n-r)^2, \quad \dim \mathcal{M} + \dim \mathcal{K} = r(n-r)
\]
and \( \mathcal{M}, \mathcal{K} \) are the full annihilators of each other with respect to the bilinear form \( \theta \). Moreover, identifying \( \mathcal{L} \) with \( \left\{ \begin{pmatrix} L & 0 \\ 0 & 0 \end{pmatrix} : L \in \mathcal{L} \right\} \) etc., we have \( \text{range}(\phi) = \mathcal{L} \oplus \mathcal{M} \oplus \mathcal{N} \).

In particular \( \text{range}(\phi) \) contains a matrix \( \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \) and so \( \mathcal{X} \) contains a matrix \( \begin{pmatrix} 0 & K \\ 0 & I \end{pmatrix} \). This latter matrix is similar to \( \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \) and so we could have taken this matrix as our matrix \( A \) to begin with. Let us suppose that this had been done. With this choice of \( A \) the bilinear form \( \theta(U, V) \) is easily seen to be \( \text{tr}(UV) \).

Now let \( X = \begin{pmatrix} H & J \\ M & N \end{pmatrix} \) be any matrix in \( \mathcal{X} \) and consider the characteristic polynomial \( \det(X + \alpha A - \lambda I) \). In this polynomial the coefficients of
\[a^{n-r}, a^{n-r+1}, \ldots, a^{n-r+1} \] are all zero and are plainly equal to the coefficients of \(X^0, X^1, \ldots, X^{n-1}\) in \(\det (H - \lambda I)\). Thus \(H\) is nilpotent. Let \(\mathcal{H}\) be the space of matrices \(H\) arising in this way. Then \(\frac{1}{2}r(r-1) = \dim \mathcal{L} \leq \dim \mathcal{H} \leq \frac{1}{2}r(r-1)\) by Gerstenhaber's theorem and so \(\dim \mathcal{H} = \frac{1}{2}r(r-1)\). By Gerstenhaber's theorem again \(\mathcal{H}\) may be reduced to the full algebra of strictly lower triangular matrices with a similarity matrix \(T\). Therefore by transforming \(\mathcal{H}\) by the similarity matrix \(\begin{pmatrix} T & 0 \\ 0 & I \end{pmatrix}\) we can take \(\mathcal{H}\) to be the full algebra of strictly lower triangular matrices. Having done this we may compute the coefficient of \(a^{n-r+1}X^{n-1}\) in \(\det (X + \alpha \lambda - \lambda I)\). This yields \(\text{tr} (JM) = 0\).

Consider any matrix \(Y = \begin{pmatrix} H & K \\ 0 & N \end{pmatrix}\) in \(\mathcal{Y}\). Then for all \(M \in \mathcal{M}\) there exists \(Z = \begin{pmatrix} 0 & J \\ M & 0 \end{pmatrix}\) in \(\mathcal{Y}\) and we have \(\text{tr} (JM) = 0\) and, since \(Y + Z \in \mathcal{Y}\), \(\text{tr} ((K + J)M) = 0\). Hence \(\text{tr} (KM) = 0\) and \(K\) is in the annihilator of \(\mathcal{M}\) with respect to \(\theta\) i.e. \(K \in \mathcal{H}\). Hence \(\mathcal{H}\) contains \(\begin{pmatrix} H & 0 \\ 0 & N \end{pmatrix}\) for all strictly lower triangular \(H\) and all \(N\).

Now we show that either every matrix in \(\mathcal{H}\) has first row zero or that every matrix in \(\mathcal{H}\) has \(r\)-th column zero. Suppose that this is not so, in which case we may find some matrix \(X = (x_{ij})\) which has neither first row nor \(r\)-th column zero, say \(x_{1,j}, x_{i,r} \neq 0\) for some \(i, j > r\). As shown above \(\mathcal{H}\) contains also the matrix \(\begin{pmatrix} H & 0 \\ 0 & N \end{pmatrix}\) where \(H\) has ones in the \((t+1, t)\) position, \(t = 1, 2, \ldots, r\), and zeros elsewhere and \(N\) is any matrix with a single one in every row and column except for the \(i\)-th row and \(j\)-th column and zeros elsewhere. But then \(x_{1,j}x_{i,r}\) is the coefficient of \(a^{n-2}\) in \(\det (X + \alpha W)\) and so is zero since \(r \geq 1\), a contradiction.

In terms of matrices the existence part of the theorem asserts that the matrices of \(\mathcal{H}\) are simultaneously similar to matrices

\[
\begin{pmatrix}
L_1 & 0 & 0 \\
A & L_2 & B \\
C & 0 & D
\end{pmatrix}
\]

where \(L_1, L_2\) are strictly lower triangular matrices of degrees \(r_1, r_2\) with \(r_1 + r_2 = r\). We can now prove this statement by induction taking as an inductive hypothesis that the statement is true for smaller values of \(n\) and all appropriate values of \(r\).

In the case where the matrices of \(\mathcal{H}\) have zero first row we consider the space of \((n-1) \times (n-1)\) matrices \(\tilde{\mathcal{H}}\) obtained by deleting the first row and first column. It is clear that each matrix of \(\tilde{\mathcal{H}}\) has at least \(r-1\) zero eigenvalues and hence \(\dim \tilde{\mathcal{H}} \leq \frac{1}{2}(r-1)(r-2) + (n-1)(n-r)\). On the other hand from the definition of \(\tilde{\mathcal{H}}\), \(\dim \tilde{\mathcal{H}} = \dim \mathcal{H} - (n-1) = \frac{1}{2}(r-1)(r-2) + (n-1)(n-r)\). Thus equality holds and the inductive hypothesis can be applied to \(\tilde{\mathcal{H}}\). Hence, for some non-singular \((n-1) \times (n-1)\) matrix \(T\) the matrices of \(T^{-1} \tilde{\mathcal{H}} T\) have the form

\[
\begin{pmatrix}
L'_1 & 0 & 0 \\
A & L_2 & B \\
C & 0 & D
\end{pmatrix}
\]
with $L_1', L_2$ strictly lower triangular of degree $r_1 - 1, r_2$ and $r_1 - 1 + r_2 = r - 1$. But then taking $S = \begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix}$ the matrices of $S^{-1} \mathcal{A} S$ have the required form.

In the case where the matrices of $\mathcal{A}$ have zero $r$-th column we use a similar argument taking $\mathcal{A}$ to be the space obtained by deleting the $r$-th row and $r$-th column from the matrices of $\mathcal{A}$.

Finally, returning to the language of vector spaces and linear mappings, suppose that $U_1, U_2$ are two subspaces of $V$ which have all the stated properties of $V_1, V_2$. Let

$$W_1 = U_1 + V_1, W_2 = U_2 \cap V_2, \dim W_1 = s_1, \dim V/W_2 = s_2, \dim W_1 \cap W_2 = t.$$ 

Then $\mathcal{A}$ induces nilpotent mappings in both $W_1$ and $V/W_2$ and therefore $\mathcal{A}, W_1, W_2$ satisfy all the conditions of the lemma. It follows that

$$\frac{1}{2}r(r - 1) + n(n - r) \leq \frac{1}{2}r(r - 1) + n(n - r) - (s_1 - t)(n - s_2 - t).$$

But $(s_1 - t)(n - s_2 - t)$ is non-negative and so is zero. Hence $s_1 = t$ or $n = s_2 + t$. From $s_1 = t$ it follows that $W_1 \leq W_2$ i.e. $U_1 + V_1 \leq U_2 \cap V_2$ and hence $U_1 = V_1, U_2 = V_2$ since otherwise all transformations in $\mathcal{A}$ would have more than $r$ zero eigenvalues, contradicting Theorem 1. From $n = s_2 + t$ it follows that $W_2 \leq W_1$ and $\mathcal{A}$ is a space of nilpotent matrices i.e. $r = n$.

References


Department of Computing Mathematics,
University College,
Cardiff.