

PRIMITIVE SPACES OF MATRICES OF BOUNDED RANK

M. D. ATKINSON and S. LLOYD

(Received 29 May 1980)

Communicated by D. E. Taylor

Abstract

A weak canonical form is derived for vector spaces of $m \times n$ matrices all of rank at most r . This shows that the structure of such spaces is controlled by the structure of an associated 'primitive' space. In the case of primitive spaces it is shown that m and n are bounded by functions of r and that these bounds are tight.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*): 15 A 30, 15 A 03.

The study of vector spaces \mathcal{X} whose vectors are $m \times n$ matrices of rank bounded by some number r was begun by Flanders (1962). He showed that such spaces necessarily have dimension at most $\max(mr, nr)$ and he classified the spaces of this maximal dimension. His work was extended by Atkinson and Lloyd (1980) to $\dim \mathcal{X} \geq \max(mr, nr) - r + 1$ while Atkinson and Stephens (1977) treated the case $\dim \mathcal{X} = 2$. In this article we shall derive a weaker classification theorem which however is valid for \mathcal{X} of arbitrary dimension. This theorem shows that the structure of \mathcal{X} depends essentially on an associated 'primitive' space with similar properties to \mathcal{X} but for which extra information is available. As an application of this result we shall deduce a result which resembles the main theorem of Atkinson and Lloyd (1980) (although it neither implies nor is implied by this theorem). Most of our methods are valid (as in the above works) only when the underlying field has at least $r + 1$ elements and this condition will be a tacit assumption in all our results.

For any space \mathcal{X} of $m \times n$ matrices we let $\rho(\mathcal{X})$ be the maximum rank of the various matrices in \mathcal{X} . If P and Q are non-singular $m \times m$ and $n \times n$ matrices

then the set

$$P\mathcal{X}Q = \{PXQ: X \in \mathcal{X}\}$$

is also a space and obviously $\rho(P\mathcal{X}Q) = \rho(\mathcal{X})$. Clearly $P\mathcal{X}Q$ is obtained from \mathcal{X} by performing a fixed sequence of row and column operations on each matrix of \mathcal{X} . In our study of the effect of $\rho(\mathcal{X})$ upon \mathcal{X} we shall freely replace \mathcal{X} by such *equivalent* spaces. If we were to regard \mathcal{X} as a space of linear mappings between vector spaces V_m and V_n (or as a space of bilinear forms on $V_m \times V_n$) such replacements would of course just correspond to changing the bases of V_m and V_n .

There are three ways in which \mathcal{X} may reduce to a ‘smaller’ space. It may be equivalent to a space of matrices all of which have a fixed row or column equal to zero, or to a space of matrices of the form (uY) where u is a column vector and $\text{rank}(Y) \leq \rho(\mathcal{X}) - 1$, or to a space of matrices of the form $(\begin{smallmatrix} v \\ Y \end{smallmatrix})$ where v is a row vector and again $\text{rank}(Y) \leq \rho(\mathcal{X}) - 1$. If any of these occur we shall say that \mathcal{X} is *imprimitive*. In the first case we clearly lose nothing by deleting the zero row or column, while in the second and third cases much of the complication within \mathcal{X} will be inherited by the space \mathcal{Y} formed by the matrices Y ; thus an understanding of the structure of \mathcal{X} may be gained by studying \mathcal{Y} which has the property $\rho(\mathcal{Y}) < \rho(\mathcal{X})$. If \mathcal{X} cannot be reduced in any of these ways it will be said to be *primitive*.

It is convenient in considering $m \times n$ matrices to allow one or both of m and n to be zero (corresponding to mappings between vector spaces one or both of which is the zero space). If we do this then the zero space consisting of a single zero by zero matrix is a primitive space; not surprisingly we call this the trivial space.

More interesting examples of primitive spaces arise out of the following construction. Let V be an $(r + 1)$ -dimensional vector space and let $\wedge^2 V$ denote its exterior square of dimension $\frac{1}{2}r(r + 1)$. Each vector $v \in V$ induces a linear mapping $x \mapsto x \wedge v$ from V into $\wedge^2 V$ which, since $v \wedge v = 0$, has rank at most r (in fact, if $v \neq 0$, it has rank exactly r). These mappings give an $(r + 1)$ -dimensional space \mathcal{W}_r of $(r + 1) \times \frac{1}{2}r(r + 1)$ matrices with $\rho(\mathcal{W}_r) = r$.

LEMMA 1. *If $r \geq 2$ then \mathcal{W}_r is primitive.*

PROOF. If the matrices of \mathcal{W}_r are equivalent to matrices (uY) with $\text{rank}(Y) \leq r - 1$ then $\wedge^2 V$ has a 1-dimensional subspace W such that for all $v \in V$ the composite mapping

$$x \mapsto x \wedge v \mapsto x \wedge v + W$$

from V into $\wedge^2 V/W$ has rank at most $r - 1$. The kernel of such a mapping must have dimension at least 2 and so, for all v , there exists \bar{v} independent of v

with $\bar{v} \wedge v \in W$. It now follows that $\dim V \leq 2$ (and so $r \leq 1$); for otherwise we could take any $v_1 \neq 0$ and any v_2 independent of $\{v_1, \bar{v}_1\}$ and there would be no vector \bar{v}_2 independent of v_2 with $\bar{v}_2 \wedge v_2 \in \langle \bar{v}_1 \wedge v_1 \rangle$. This disposes of one of the ways in which \mathfrak{W}_r could be imprimitive; the others are easier to deal with and we leave the routine checks to the reader.

\mathfrak{W}_1 and, more generally, any space \mathfrak{X} with $\rho(\mathfrak{X}) = 1$ is imprimitive; this is a consequence of the following observation due to Westwick (1972), Theorem 4.2, and which can be deduced also from Lemmas 4 and 5 below.

LEMMA 2. *A space of matrices \mathfrak{X} with $\rho(\mathfrak{X}) = 1$ is equivalent either to a space of matrices with non-zero entries in the first row only, or to a space of matrices with entries in the first column only.*

More examples of primitive spaces can be derived from the higher exterior powers of an n -dimensional vector space V . For each non-zero $v \in V$ the mapping $x \mapsto x \wedge v$ from $\wedge^k V$ into $\wedge^{k+1} V$ has rank $\binom{n-1}{k}$. A similar proof to that of Lemma 1 shows that the corresponding n -dimensional space of $\binom{n}{k} \times \binom{n}{k+1}$ matrices is primitive if $1 \leq k \leq n - 2$. Another source of primitive spaces is provided by the next lemma whose proof follows directly from the definitions.

LEMMA 3. *If \mathfrak{X}_1 and \mathfrak{X}_2 are primitive spaces then the space $\mathfrak{X}_1 \oplus \mathfrak{X}_2$ of all matrices*

$$\begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \text{ with } X_1 \in \mathfrak{X}_1 \text{ and } X_2 \in \mathfrak{X}_2$$

is also primitive.

Our main result is

THEOREM 1. *If \mathfrak{X} is a space of $m \times n$ matrices then there exists a primitive space \mathfrak{Y} and integers $p, q \geq 0$ with $\rho(\mathfrak{X}) = p + q + \rho(\mathfrak{Y})$ such that \mathfrak{X} is equivalent to a space of matrices of the form*

$$\left[\begin{array}{c|cc} p \times q & & \\ \hline & Y & 0 \\ & 0 & 0 \end{array} \right]$$

where the submatrices Y constitute the primitive space \mathfrak{Y} .

Moreover for a primitive space \mathfrak{X} ($p = 0, q = 0, \mathfrak{X} = \mathfrak{Y}$) with $r = \rho(\mathfrak{X})$ one of the following occurs:

- (i) $m = r + 1, n \leq \frac{1}{2}r(r + 1)$,

- (ii) $m \leq \frac{1}{2}r(r + 1), n = r + 1,$
- (iii) *for some integers $c, d \geq 2$ with $c + d = r, m \leq c + 1 + \frac{1}{2}d(d + 1), n \leq d + 1 + \frac{1}{2}c(c + 1).$*

The bounds in (i), (ii) and (iii) are all optimal. This is demonstrated in case (i) by \mathcal{W}_r while its transpose \mathcal{W}'_r deals with case (ii); for case (iii) we rely on Lemma 3 which shows that $\mathcal{W}'_c \oplus \mathcal{W}'_d$ is primitive.

The primitive space \mathcal{Y} in the theorem is, in general, not uniquely determined by \mathcal{X} . To see an example of this observe that the 4×6 matrices of \mathcal{W}_3 are equivalent to matrices

$$\left(\begin{array}{c|c} & B \\ \hline A & \text{000} \end{array} \right)$$

where the space of all the 3×3 matrices B is \mathcal{W}_2 . Then the space \mathcal{X} of all 6×6 matrices

$$\left(\begin{array}{c|c} A & \begin{array}{c} B \\ \text{000} \end{array} \\ \hline C & \begin{array}{c} \text{000} \\ \text{000} \end{array} \end{array} \right)$$

where C ranges over all 2×3 matrices has $\rho(\mathcal{X}) = 5$ and satisfies the theorem with $p = 0, q = 3, \mathcal{Y} = \mathcal{W}_2$. On the other hand simple row operations show that it also satisfies the theorem with $p = 2, q = 0, \mathcal{Y} = \mathcal{W}_3$.

We note that the theorem applies in the special case, considered by Westwick (1972), of a space all of whose non-zero matrices have the same rank; in this case the space \mathcal{Y} will also inherit this extra property.

Among the technical lemmas required to prove the theorem are two previously published results which for ease of reference we restate here.

LEMMA 4 (Flanders (1962)). *If \mathcal{X} is a space of matrices with $\rho(\mathcal{X}) = r$ then there is a space equivalent to \mathcal{X} whose matrices have the form $\begin{pmatrix} T & U \\ V & 0 \end{pmatrix}$ where T is an $r \times r$ matrix.*

LEMMA 5 (The proof of Lemma 4, Atkinson and Lloyd (1980)). *Let \mathcal{X} be a space of matrices all partitioned in some fixed way as $\begin{pmatrix} W & U \\ V & 0 \end{pmatrix}$ and let \mathcal{U}, \mathcal{V} be the spaces formed by the submatrices U, V respectively. Then $\rho(\mathcal{U}) + \rho(\mathcal{V}) \leq \rho(\mathcal{X})$.*

DEFINITION. A space \mathcal{X} of matrices is said to have the *row condition* if, whenever \mathcal{X} is equivalent to a space of matrices all partitioned in some fixed way as $\begin{pmatrix} A & B \\ C & 0 \end{pmatrix}$, all the submatrices B have rank less than the number of rows of B .

Similarly \mathcal{X} is said to have the *column condition* if, whenever \mathcal{X} is equivalent to a space of matrices partitioned in this way, all the submatrices C have rank less than the number of columns of C .

LEMMA 6. *If a space \mathcal{X} with $\rho(\mathcal{X}) = r$ has the row condition then there is a space equivalent to \mathcal{X} whose matrices have non-zero entries only in their first $\frac{1}{2}r(r+1)$ columns. Moreover, either*

(i) *these matrices have non-zero entries in their first $1 + \frac{1}{2}r(r-1)$ columns only, or*

(ii) *these matrices have non-zero entries in their first $r+1$ rows only.*

PROOF. If $\rho(\mathcal{X}) = 0$ the lemma is obviously correct and if $\rho(\mathcal{X}) = 1$ it follows from Lemma 2. These cases provide the first steps of an induction. We shall assume now that $r = \rho(\mathcal{X}) \geq 2$ and, as an inductive hypothesis, that the lemma is true for spaces of matrices of rank smaller than r . By using Lemma 4 we may replace \mathcal{X} by a space whose matrices all have the form $\begin{pmatrix} W & Y \\ Z & 0 \end{pmatrix}$ for some $r \times r$ matrices W . Let \mathcal{Y} be the space formed by the submatrices Y and \mathcal{Z} the space formed by the submatrices Z . Then Y also satisfies the row condition. For if there were row and column operations which reduced \mathcal{Y} to a form which contravened the definition of the row condition then these operations performed on \mathcal{X} would show that \mathcal{X} did not satisfy the row condition.

In particular, since each $Y \in \mathcal{Y}$ has r rows, $\rho(\mathcal{Y}) < r-1$. Hence, by the inductive hypothesis, \mathcal{Y} must be equivalent to a space whose matrices have non-zero entries in their first $\frac{1}{2}r(r-1)$ columns only, and so \mathcal{X} is equivalent to a space whose matrices have non-zero entries in their first $r + \frac{1}{2}r(r-1) = \frac{1}{2}r(r+1)$ columns only. This establishes the first part of the lemma.

If, in fact, $\rho(\mathcal{Y}) < r-1$ then, by the same reasoning, we would obtain conclusion (i) of the second part since $r + \frac{1}{2}(r-1)(r-2) = 1 + \frac{1}{2}r(r-1)$. Consequently we may as well assume that $\rho(\mathcal{Y}) = r-1$ and hence, by Lemma 5, $\rho(\mathcal{Z}) < 1$ and we can apply Lemma 2 to \mathcal{Z} . If \mathcal{Z} is equivalent to a space of matrices with zeros everywhere except in their first row we obtain conclusion (ii) of the lemma and so we may assume that \mathcal{Z} can, and has been, replaced by matrices with non-zero entries in their first column only. Thus \mathcal{X} consists of matrices

$$\begin{pmatrix} \vdots & U \\ \vdots & 0 \end{pmatrix}$$

where each U has r rows. As above the space \mathcal{U} of all submatrices U satisfies the row condition and has $\rho(\mathcal{U}) < r-1$. A final application of the inductive hypothesis (to \mathcal{U}) shows that \mathcal{X} is equivalent to a space of matrices with

non-zero entries in their first $1 + \frac{1}{2}r(r - 1)$ columns only, that is conclusion (i) holds.

Of course, the same result holds with the words ‘row’ and ‘column’ interchanged.

PROOF OF THEOREM 1. The first part of the theorem obviously holds when $m = n = 0$ (\mathcal{X} is then the trivial space) and this provides the base of an easy induction on $m + n$. In carrying out the inductive step we may take \mathcal{X} to be imprimitive since in the primitive case we can set $p = q = 0$ and $\mathcal{Y} = \mathcal{X}$. If \mathcal{X} is equivalent to a space of matrices (uU) for column vectors u and $m \times (n - 1)$ matrices U of rank at most $\rho(\mathcal{X}) - 1$ we apply the inductive hypothesis to the space \mathcal{U} formed by the matrices U . Then \mathcal{U} is equivalent to a space of matrices of the form

$$\left(\begin{array}{c|cc} p \times q & & \\ \hline & Y & 0 \\ & 0 & 0 \end{array} \right)$$

with \mathcal{Y} primitive and $p + q + \rho(\mathcal{Y}) = \rho(\mathcal{X}) - 1$. Thus \mathcal{X} is equivalent to a space of matrices of the form

$$\left(\begin{array}{c|cc} p \times (q + 1) & & \\ \hline & Y & 0 \\ & 0 & 0 \end{array} \right)$$

which have the form claimed in the first part of the theorem. Similar arguments deal with the remaining two ways in which \mathcal{X} can be imprimitive.

For the second part (with \mathcal{X} primitive) we begin by replacing \mathcal{X} by a space of matrices of the form

$$\left(\begin{array}{c|c} a \times b & U \\ \hline V & 0 \end{array} \right)$$

for some fixed a, b with $a \leq r, b \leq r$. By Lemma 4 such equivalent spaces definitely exist, and we shall choose one with $a + b$ as small as possible. Let \mathcal{U}, \mathcal{V} be the spaces formed by the submatrices U, V respectively. Then \mathcal{U} satisfies the row condition. For, if on the contrary, for some $k > 0$ the matrices of \mathcal{U} are equivalent to matrices

$$\left(\begin{array}{c|c} & U_1 \\ \hline U_2 & 0 \end{array} \right)$$

where each U_1 has k rows and some U_1 has rank k then the matrices of \mathcal{X} are equivalent to matrices

$$\left(\begin{array}{c|c} & U_1 \\ \hline W & 0 \end{array} \right)$$

where, by Lemma 5, the space \mathcal{W} of submatrices W satisfies $\rho(\mathcal{W}) \leq r - k$. But then, repartitioning these matrices equivalent to \mathcal{X} as $\binom{e}{Y}$ we have, for all Y , $\text{rank}(Y) \leq r - k + k - 1 = r - 1$ contradicting the primitivity of \mathcal{X} . For precisely similar reasons \mathcal{V} satisfies the column condition. Moreover, if $x = \rho(\mathcal{U})$ and $y = \rho(\mathcal{V})$ then $x + y \leq r$ and certainly, since \mathcal{X} is primitive, $x < r$ and $y < r$.

If $m = r + 1$ then we can obtain possibility (i) since $n \leq b + \frac{1}{2}x(x + 1) < r + \frac{1}{2}r(r - 1) = \frac{1}{2}r(r + 1)$. So in the remainder of the proof we may take $m > r + 1$ and likewise $n > r + 1$ and work towards obtaining possibility (iii).

Now we apply Lemma 6 to \mathcal{U} . If conclusion (i) of this lemma holds then

$$(1) \quad n \leq b + \frac{1}{2}x(x - 1) + 1$$

while if conclusion (ii) holds then

$$(1') \quad n \leq b + \frac{1}{2}x(x + 1) \quad \text{and} \quad a = x + 1$$

(the last equation following from the minimality of $a + b$). Similarly, arguing with \mathcal{V} , we have either

$$(2) \quad m \leq a + \frac{1}{2}y(y - 1) + 1,$$

or

$$(2') \quad m \leq a + \frac{1}{2}y(y + 1) \quad \text{and} \quad b = y + 1.$$

If (1) and (2) hold we may satisfy the theorem by choosing any c, d with $x < c, y < d$ and $c + d = r$. For then

$$m \leq r + \frac{1}{2}y(y - 1) + 1 \leq c + d + \frac{1}{2}d(d - 1) + 1 = c + 1 + \frac{1}{2}d(d + 1)$$

and similarly $n \leq d + 1 + \frac{1}{2}c(c + 1)$.

If (1) and (2') hold we put $d = y + 1$ and $c = r - d \geq x + y - y - 1 = x - 1$. Then

$$\begin{aligned} m &\leq r + \frac{1}{2}y(y + 1) = c + d + \frac{1}{2}d(d - 1) \\ &< c + 1 + \frac{1}{2}d(d + 1) \end{aligned}$$

and

$$\begin{aligned} n &\leq y + 1 + \frac{1}{2}x(x - 1) + 1 \\ &\leq d + 1 + \frac{1}{2}c(c + 1). \end{aligned}$$

If (1') and (2) hold a similar argument can be used while if (1') and (2') hold we obtain the conclusion by again choosing any c, d with $x < c, y < d$ and $c + d = r$.

Finally we observe that the values c, d obtained are necessarily each greater than 1 since we have assumed $m, n > r + 1$.

In Atkinson and Lloyd (1980) a space \mathcal{X} of matrices with $\rho(\mathcal{X}) = r$ was defined to be r -decomposable if it was equivalent to a space of matrices all of the form

$$\left(\begin{array}{c|c} p \times q & \\ \hline & 0 \end{array} \right)$$

for some fixed p, q with $p + q = r$. In the terminology of Theorem 1, r -decomposable spaces are those which satisfy the theorem with \mathcal{Y} equal to the trivial space. For all m, n and $r \geq 2$ the space of all matrices

$$\left(\begin{array}{cc} Z & \\ Y & 0 \\ 0 & 0 \end{array} \right)$$

(Z an arbitrary $(r - 2) \times n$ matrix and $Y \in \mathcal{W}_2$) is an example of a space \mathcal{X} with $\rho(\mathcal{X}) = r$ of 'large' dimension $n(r - 2) + 3$ which is not r -decomposable. To complement this example we have

THEOREM 2. *If \mathcal{X} is a space of $m \times n$ matrices with $m < n$, $\rho(\mathcal{X}) = r$, $n > \frac{1}{2}r^2(r + 1)/(r - 2)$ and $\dim \mathcal{X} > n(r - 2) + 3$ then \mathcal{X} is r -decomposable.*

PROOF. We use the notation of Theorem 1 and take \mathcal{X} in the canonical form given in that theorem. Then

$$\begin{aligned} n(r - 2) + 3 < \dim \mathcal{X} &\leq mq + np - pq + \dim \mathcal{Y} \\ &\leq n(p + q) + \dim \mathcal{Y} \end{aligned}$$

and putting $s = \rho(\mathcal{Y}) = r - p - q$ we obtain

$$(3) \quad n(s - 2) + 3 < \dim \mathcal{Y}.$$

By Theorem 1 the maximum number of rows or columns the matrices of \mathcal{Y} can have is $\frac{1}{2}s(s + 1)$ and so from Theorem 1 of Flanders (1962) $n(s - 2) + 3 < \frac{1}{2}s^2(s + 1)$. But if $s > 2$ we would have

$$\frac{1}{2}r^2(r + 1)/(r - 2) < \frac{1}{2}s^2(s + 1)/(s - 2)$$

which, since $s \leq r$, is impossible; hence $s = 0, 1$ or 2 . By Lemma 2, $s \neq 1$ and we just have to exclude the case $s = 2$. In this case (3) shows that $\dim \mathcal{Y} > 3$ which contradicts

LEMMA 7. *Every primitive space \mathcal{X} with $\rho(\mathcal{X}) = 2$ is a 3-dimensional space of 3×3 matrices.*

PROOF. If \mathcal{X} consists of $m \times n$ matrices then $m, n \geq 3$ from the definition of primitivity. On the other hand Theorem 1 shows that $m, n < 3$. If \mathcal{X} contains

no matrix of rank 1 then Theorem 5.2 of Westwick (1972) proves that \mathcal{X} is 3-dimensional. To complete the proof therefore it is sufficient to show that a space \mathcal{X} of 3×3 matrices with $\rho(\mathcal{X}) = 2$ and containing a rank 1 matrix M is 2-decomposable (and thus imprimitive).

By passing to an equivalent space we may take

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

in which case, by Lemma 2 of Atkinson and Lloyd (1980), the matrices of \mathcal{X} have the form

$$\begin{pmatrix} * & * & * \\ * & & Y \\ * & & \end{pmatrix}$$

where $\text{rank } Y \leq 1$. Using Lemma 2 on the space formed by the submatrices Y and replacing \mathcal{X} by an equivalent space the matrices may be assumed to all have the form

$$\begin{pmatrix} * & * & * \\ * & * & 0 \\ * & * & 0 \end{pmatrix}$$

or to all have the form

$$\begin{pmatrix} * & * & * \\ * & * & * \\ * & 0 & 0 \end{pmatrix}.$$

To within transposes these cases are the same and so we lose no generality in assuming the former. Then, by Lemma 5, the third column of each matrix is zero (so that \mathcal{X} is 2-decomposable) or the matrices have the form

$$\begin{pmatrix} * & * & * \\ & Z & 0 \\ & & 0 \end{pmatrix}$$

with $\text{rank } Z \leq 1$.

In this last case Lemma 2 may be applied to the space formed by the submatrices Z . It shows that \mathcal{X} is equivalent to a space of matrices of the form

$$\begin{pmatrix} * & * & * \\ * & 0 & 0 \\ * & 0 & 0 \end{pmatrix}$$

or to a space of matrices of the form

$$\begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & 0 \end{pmatrix},$$

both of which are 2-decomposable, completing the proof.

By a rather longer calculation Lemma 7 can be extended to show that $\mathcal{O}\mathbb{S}_2$ is the only primitive space with $\rho(\mathcal{X}) = 2$. Much more extensive computations have been performed by one of us (Lloyd) resulting in the classification of primitive spaces with $\rho(\mathcal{X}) = 3$.

References

- M. D. Atkinson and S. Lloyd (1980), 'Large spaces of matrices of bounded rank', *Quart. J. Math. Oxford Ser.* **31**, 253–262.
- M. D. Atkinson and N. M. Stephens (1978), 'Spaces of matrices of bounded rank', *Quart. J. Math. Oxford Ser.* **28**, 205–207.
- H. Flanders (1962), 'On spaces of linear transformations of bounded rank', *J. London Math. Soc.* **37**, 10–16.
- R. Westwick (1972), 'Spaces of linear transformations of equal rank', *Linear Algebra and Appl.* **5**, 49–64.

Department of Computing Mathematics
University College
Cardiff, Wales