

## THE ROW SPACE OF A MATRIX<sup>1</sup>

The row space  $\mathcal{R}(A)$  of a matrix  $A$  is defined as the vector space spanned by the rows of  $A$ . If  $B$  is a matrix with the same number of columns as the number of rows of  $A$ , then  $B$  and  $A$  can be multiplied (in this order, i.e., the matrix  $BA$  is defined). It is easily seen by the rule two matrices are multiplied that the rows of  $BA$  are linear combinations of the rows of  $A$  (where the coefficients for the linear combination are supplied by the row in  $B$  having the same index as the row of  $BA$  is question). Therefore, each row of  $BA$  is in the space  $\mathcal{R}(A)$ , and so  $\mathcal{R}(BA) \subset \mathcal{R}(A)$ . If  $B$  is an invertible square matrix, then  $\mathcal{R}(A) = \mathcal{R}((B^{-1}B)A) \subset \mathcal{R}(B^{-1}(BA)) \subset \mathcal{R}(BA)$  also holds, and so we also have  $\mathcal{R}(BA) = \mathcal{R}(A)$  in this case.

There are a few important consequences of these remarks. First, if  $C$  is the row echelon form of  $A$ , then  $\mathcal{R}(C) = \mathcal{R}(A)$ ; this is because  $C = PA$  for an invertible matrix  $P$  (namely,  $P$  is the product of the elementary matrices used in obtaining the row echelon form of  $A$ ). Therefore, writing  $\dim(V)$  for the dimension of the vector space  $V$ ,  $\dim(\mathcal{R}(A)) = \dim(\mathcal{R}(C))$ .<sup>2</sup> The latter equals the number of nonzero rows in  $C$ ; this is because these rows are linearly independent (since each has a leading entry in a column where all the other rows are zero), so they form a basis for the space they span. This number is also the same as the number of leading entries of  $C$ , because in the row echelon form, each nonzero row has exactly one leading entry.

For an integer  $k \geq 0$ , let  $A \upharpoonright k$  be the matrix that is obtained by keeping only those entries of  $A$  that are located in the first  $k$  columns.<sup>3</sup> The resulting matrix will have  $k$  or fewer columns.<sup>4</sup> A few observations here will be useful. First, for any integer  $k \geq 0$  and any matrix  $A$  we have<sup>5</sup>

$$\mathcal{R}(A \upharpoonright k) = \{v \upharpoonright k : v \in \mathcal{R}(A)\};$$

the right-hand side here will be written as  $\mathcal{R}(A) \upharpoonright k$ .<sup>6</sup> This is because the left-hand side is the set of all linear combinations of the vectors  $a_{i*} \upharpoonright k$ , where  $a_{i*}$  denotes the  $i$ th row of  $A$ . Therefore, if  $\mathcal{R}(A) = \mathcal{R}(B)$  for the matrices  $A$  and  $B$ , then  $\mathcal{R}(A \upharpoonright k) = \mathcal{R}(B \upharpoonright k)$  for any integer  $k \geq 0$ . After these comments we are ready to prove the following.

**Theorem.** *Let  $A$  and  $B$  be matrices of the same size such that  $\mathcal{R}(A) = \mathcal{R}(B)$ . Then the row echelon forms of  $A$  and  $B$  are the same.*

An important consequence of this is that the row echelon form of a matrix is unique. Namely, if  $A = B$ , then certainly  $\mathcal{R}(A) = \mathcal{R}(B)$ . Even though we talked about *the* row echelon form of a matrix rather than *a* row echelon form, we did this only in anticipation of this result.

*Proof.* Let  $C$  and  $D$  be row echelon forms of  $A$  and  $B$ , respectively; we will show that  $C = D$ . Let  $n$  be the number of columns of  $A$  (and of  $B$ , since they have the same size). We will use induction on  $n$ ; the result is clearly true for  $n = 0$ .<sup>7</sup>

<sup>1</sup>Notes for Course Mathematics 10.1 at Brooklyn College of CUNY. Attila Máté, October 5, 2007.

<sup>2</sup>Because  $\mathcal{R}(A)$  and  $\mathcal{R}(C)$  are the same vector space, according to the what was said above.

<sup>3</sup>The main use of the symbol  $\upharpoonright$ , called restriction, in mathematics is  $f \upharpoonright A$  for a function  $f$  and a subset  $A$  of its domain. This denotes a function  $g$  such that  $g(x) = f(x)$  whenever  $x \in A$ , and  $g$  is undefined if  $x \notin A$ . The function  $f \upharpoonright A$  is naturally called the restriction of  $f$  to  $A$ . Here we use this symbol with a different meaning.

<sup>4</sup>Fewer than  $k$  if  $A$  had fewer than  $k$  columns to begin with, and in this case  $A \upharpoonright k = A$ ; otherwise  $A$  will have exactly  $k$  columns. The case  $k = 0$  seems strange, but it will simplify things in the proof below; of course  $A \upharpoonright 0$  will be an empty matrix (a matrix with no entries).

<sup>5</sup>The notation  $v \upharpoonright k$  used in the next line is to be interpreted by taking the row vector  $v$  to be a matrix with one row.

<sup>6</sup>The notation  $\mathcal{R}(A) \upharpoonright k$  introduced after the displayed line is a natural extension of notation introduced above. Similarly, for any set  $V$  of row vectors one can write  $V \upharpoonright k = \{v \upharpoonright k : v \in V\}$ .

<sup>7</sup>Because then  $C$  and  $D$  will both be the empty matrices, and any two empty matrices are equal. Or, more clearly, we will consider them equal – because there is no way to distinguish them, except perhaps by the number of rows. If one takes the view that these matrices have  $m$  rows and 0 columns, then one would consider two empty matrices different if they have different

Now, let  $n \geq 1$ , and assume that the result is true for any two matrices with  $n - 1$  columns. It is clear that a row echelon form of  $A \uparrow (n - 1)$  is  $C \uparrow (n - 1)$ ; this is because using the same elementary row operations for  $A \uparrow (n - 1)$  that were used in obtaining  $C$  from  $A$  will result in the matrix  $C \uparrow (n - 1)$ , and, clearly,  $C \uparrow (n - 1)$  is in row echelon form.<sup>8</sup> Similarly, a row echelon form of  $B \uparrow (n - 1)$  is  $D \uparrow (n - 1)$ . We have  $\mathcal{R}(A \uparrow (n - 1)) = \mathcal{R}(A) \uparrow (n - 1) = \mathcal{R}(B) \uparrow (n - 1) = \mathcal{R}(B \uparrow (n - 1))$ , by the observations above. Therefore, by the induction hypothesis, we have  $C \uparrow (n - 1) = D \uparrow (n - 1)$ .

Our first goal is to show that the leading entries of  $C$  and  $D$  are in the same columns. Since the first  $n - 1$  columns of  $C$  and  $D$  are the same, the only question is whether the last column of  $C$  or  $D$  contains a leading entry. Now, if  $\dim(\mathcal{R}(A)) = \dim((A \uparrow (n - 1)))$ , then  $C$  has the same number of leading entries as  $C \uparrow (n - 1)$ , so the last column of  $C$  cannot contain a leading entry, and if  $\dim(\mathcal{R}(A)) > \dim((A \uparrow (n - 1)))$ , then  $C$  has more leading entries than  $C \uparrow (n - 1)$ , and so the last column of  $C$  must contain a leading entry.<sup>9</sup> As  $\mathcal{R}(B) = \mathcal{R}(A)$  and  $\mathcal{R}(B) \uparrow (n - 1) = \mathcal{R}(A) \uparrow (n - 1)$ , the same considerations with the matrix  $B$  show that the last columns of  $C$  and  $D$  either both have leading entries or neither of them do. Thus,  $C$  and  $D$  indeed have leading entries in the same columns.

It will be easy to establish from this that  $C = D$ . To this end, note that the leading entries of  $C$  and  $D$  are in the same places in  $C$  and  $D$  (that is, they are also in the same rows).<sup>10</sup> Let  $r$  be the number of nonzero rows of  $C$ , and let  $l(i)$  be the column index of the leading entry in row  $i$  of  $C$  ( $1 \leq i \leq r$ ). Since  $\mathcal{R}(C) = \mathcal{R}(A) = \mathcal{R}(B) = \mathcal{R}(D)$ , every nonzero row of  $D$  is a linear combination of the nonzero rows of  $C$ . That is, writing  $C = (c_{ij})$  and  $D = (d_{ij})$ , and recalling that  $d_{i*}$  then denotes  $i$ th row of  $D$  and  $c_{j*}$ , the  $j$ th row of  $C$ , for each  $j$  with  $1 \leq j \leq r$  we have

$$(1) \quad d_{i*} = \sum_{j=1}^r \alpha_j c_{j*};$$

of course, the scalars  $\alpha_j$  depend on  $i$  as well here, but  $i$  will remain fixed in the forthcoming argument, so this dependence is not indicated. This equation stated for the entries in the  $l(k)$ th column for a  $k$  with  $1 \leq k \leq r$  says that

$$(2) \quad d_{i \ l(k)} = \sum_{j=1}^r \alpha_j c_{j \ l(k)}.$$

Using the notation

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

called Kronecker  $\delta$  after the 19th century German mathematician Leopold Kronecker, the proof can now be completed quickly. Noting that  $d_{i \ l(k)} = \delta_{ik}$  and  $c_{j \ l(k)} = \delta_{jk}$ ,<sup>11</sup> equation (2) can be written as

$$\delta_{ik} = \sum_{j=1}^r \alpha_j \delta_{jk} = \alpha_k$$

---

numbers of rows. Since we are dealing with two matrices of the same size,  $C$  and  $D$  will have the same number of rows, so we will still have  $C = D$ . It is not really worth discussing the question what the row space of the empty matrix is – but if one insists, one can take this to be the space consisting of the single row vector of length zero, this being the zero element of the space.

On the face of it, one may consider it silly to start the induction with  $n = 0$ , but this is the elegant way to do it. If one wants to start with  $n = 1$ , then one needs to include a proof for  $n = 1$ . This is simple enough to do, but it is unnecessary. After a careful reading of the proof here, it is clear that the induction step works also for the case stepping from  $n = 0$  to  $n = 1$ , and it does not even matter what one thinks of matrices with zero columns. A further remark on this issue will be made right after the proof.

<sup>8</sup>Another way of putting this argument is as follows: if  $P$  is a product of elementary matrices such that  $C = PA$  then  $C \uparrow (n - 1) = P(A \uparrow (n - 1))$ ;  $C = PA$  implies the latter equation for any matrices  $C$ ,  $P$ , and  $A$  because of the way operations can be done with block matrices; that is, consider  $A$  as the block matrix  $(A \uparrow (n - 1) \ a_{*n})$ , where  $a_{*n}$  is the last column of  $A$ .

<sup>9</sup>Because, as we pointed this out above, the number of leading entries in a row echelon form of a matrix is the same as the dimension of its row space. These considerations also imply that the only values that are possible for  $\dim(\mathcal{R}(A))$  are  $\dim(\mathcal{R}(A \uparrow (n - 1)))$  and  $\dim(\mathcal{R}(A \uparrow (n - 1))) + 1$ , since either  $C$  has the same number of leading entries as  $C \uparrow (n - 1)$  or it has one more.

<sup>10</sup>This is because each nonzero row has exactly one leading entry, and the column indices of the leading entries are an increasing function of the row indices of the nonzero rows.

<sup>11</sup>These equations express the fact that in a column containing a leading entry, the leading entry itself is 1, and every other entry is 0.

for the given  $i$  and for any  $k$  with  $1 \leq k \leq r$ ; stating this with  $j$  instead of  $k$ , we have  $\alpha_j = \delta_{ij}$  for any  $j$  with  $1 \leq j \leq r$ . Substituting this into equation (1), we obtain

$$d_{i*} = \sum_{j=1}^r \delta_{ij} c_{j*} = c_{i*}.$$

That is, the  $i$ th row of  $C$  is the same as the  $i$ th row of  $D$ . This being true for any  $i$  with  $1 \leq i \leq r$ , each nonzero row of  $D$  agrees with the corresponding row of  $C$ . Hence  $D = C$ , as we wanted to show.

*Remark.* If one wants to avoid the case  $n = 0$ , one needs to establish the result for  $n = 1$  separately. This is easy to do by observing that, in case  $n = 1$ , i.e., when  $A$  and  $B$  have only one column each, there are two cases: 1) all entries of  $A$  and  $B$  are zero, in which case  $C$  and  $D$  equal the zero matrix; 2)  $A$  and  $B$  each have nonzero entries, in which case  $C = D$  is a column vector whose first entry is 1 and all other entries are 0. It is worth noting that this is essentially the same argument as the one used in the induction step to decide whether or not the  $n$ th columns of  $C$  and  $D$  have leading entries. That is, starting the induction at  $n = 1$  instead of  $n = 0$  unnecessarily repeats a part of the argument used in the induction step.