Lecture 14. Matrix Rank
May 2023

Row Space and Column Space

The following definition defines two important vector spaces associated with a matrix.

DEFINITION 14.1 If A is an $m \times n$ matrix, then the subspace of $\mathbb{R}^n$ spanned by the row vectors of A is called the **row space** of A, and the subspace of $\mathbb{R}^m$ spanned by the column vectors of A is called the **column space** of A. We will denote the row space of A and the column space of A by $\mathbf{row}\, A$ and $\mathbf{col}\, A$, respectively.

DEFINITION 14.2

- The **row rank** of A is the dimension of the row space of A, $\dim(\mathbf{row}\, A)$.
- The **column rank** of A is the dimension of the column space of A, $\dim(\mathbf{col}\, A)$.
- The **rank** of matrix A is the number of leading 1's in any row-echelon matrix to which A can be carried by row operations. It is denoted $\operatorname{rank} A$.

Our first objective is to show that the matrix rank is well-defined and is equal to the row rank and the column rank.

Row Rank

LEMMA 14.3 Elementary row operations do not change the row space of a matrix.

Proof. It is enough to prove this statement in the case when A is transformed to B by a single row operation. Suppose

$$A = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{pmatrix},$$

where $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ denote the rows of A. The row operation $A \to B$ either interchanges two rows, multiplies a row by a nonzero constant, or adds a multiple of a row to a different row. The analysis of the first two cases is trivial. In the last case, suppose that λ times row j is added to row k where $j < k$. Then

$$B = \begin{pmatrix} \vdots \\ \mathbf{r}_j \\ \vdots \\ \mathbf{r}_k + \lambda \mathbf{r}_j \\ \vdots \end{pmatrix}.$$

Fundamental Theorem 13.7 shows that

$$\mathbf{span}\{\ldots, \mathbf{r}_j, \ldots, \mathbf{r}_k, \ldots\} = \mathbf{span}\{\ldots, \mathbf{r}_j, \ldots, \mathbf{r}_k + \lambda \mathbf{r}_j, \ldots\}.$$

That is, $\mathbf{row}\, A = \mathbf{row}\, B$. $\square$

Different series of elementary row operations can carry the same matrix A to different row-echelon matrices. However, the number of leading 1's must be the same in each of these row-echelon matrices. Hence, this number depends only on A and not on the way in which A is carried to row-echelon form. So the following result is of interest.

LEMMA 14.4 If a matrix R is in row echelon form, then the row vectors with the leading 1's (the nonzero row vectors) form a basis for the row space of R.

Proof. We first prove that the rows of R are independent. Suppose the leading 1's are located in positions $(1, j_1), (2, j_2), \dots, (k, j_k)$, where $j_1 > j_2 > \dots > j_k$. Consider the vector equation

$$\lambda_1 \mathbf{r}_1 + \lambda_2 \mathbf{r}_2 + \dots + \lambda_k \mathbf{r}_k = \mathbf{0}. \tag{1}$$

Equating the entries in position j_1 implies

$$\lambda_1 \cdot 1 + \lambda_2 \cdot 0 + \lambda_3 \cdot 0 + \dots \lambda_k \cdot 0 = 0,$$

whence it follows that $\lambda_1 = 0$. Next we can equate the entries in position j_2:

$$\lambda_2 \cdot 1 + \lambda_3 \cdot 0 + \dots \lambda_k \cdot 0 = 0,$$

whence it follows that $\lambda_2 = 0$. Now the same argument shows that $\lambda_3 = \dots = \lambda_k = 0$. This shows that (1) has only trivial solution $\lambda_1 = \dots = \lambda_k = 0$. So the rows of R are independent. They span **row** R by definition. □

Column Rank

THEOREM 14.5 Elementary row operations do not alter dependence relationships among the column vectors.

Proof. Earlier we have shown that performing elementary row operations on a matrix A can be represented as multiplying this matrix on the left by an invertible matrix E, that is, A is transformed to EA.

Suppose first that $\mathbf{c}_{j_1}, \mathbf{c}_{j_2}, \dots, \mathbf{c}_{j_k}$ are linearly dependent column vectors of A, so there are scalars $\lambda_1, \lambda_2, \dots, \lambda_k$ that are not all zero and such that

$$\lambda_1 \mathbf{c}_{j_1} + \lambda_2 \mathbf{c}_{j_2} + \dots + \lambda_k \mathbf{c}_{j_k} = \mathbf{0}. \tag{2}$$

If we perform an elementary row operation on A, then these vectors will be changed into new column vectors $E\mathbf{c}_{j_1}, E\mathbf{c}_{j_2}, \dots, E\mathbf{c}_{j_k}$. These new column vectors are linearly dependent and, in fact, related by an equation

$$\lambda_1 E\mathbf{c}_{j_1} + \lambda_2 E\mathbf{c}_{j_2} + \dots + \lambda_k E\mathbf{c}_{j_r} = E(\lambda_1 \mathbf{c}_{j_1} + \lambda_2 \mathbf{c}_{j_2} + \dots + \lambda_k \mathbf{c}_{j_k}) = \mathbf{0}$$

with exactly the same coefficients as in (2).

Conversely, if columns $E\mathbf{c}_{j_1}, E\mathbf{c}_{j_2}, \dots, E\mathbf{c}_{j_k}$ are linearly dependent, i.e.,

$$\lambda_1 E\mathbf{c}_{j_1} + \lambda_2 E\mathbf{c}_{j_2} + \dots + \lambda_k E\mathbf{c}_{j_r} = \mathbf{0}$$

is satisfied by nonzero scalars $\lambda_1, \lambda_2, \dots, \lambda_k$, then

$$\lambda_1 \mathbf{c}_{j_1} + \lambda_2 \mathbf{c}_{j_2} + \dots + \lambda_k \mathbf{c}_{j_k} = E^{-1}(\lambda_1 E\mathbf{c}_{j_1} + \lambda_2 E\mathbf{c}_{j_2} + \dots + \lambda_k E\mathbf{c}_{j_r}) = \mathbf{0}. \quad \square$$

COROLLARY 14.6 If A and B are row equivalent matrices, then:

(a) A given set of column vectors of A is linearly independent if and only if the corresponding column vectors of B are linearly independent.
(b) A given set of column vectors of A forms a basis for the column space of A if and only if the corresponding column vectors of B form a basis for the column space of B.

LEMMA 14.7 If a matrix R is in row echelon form, then the column vectors with the leading 1's of the row vectors form a basis for the column space of R.

Proof. Let $S = \{\mathbf{c}_{j_1}, \mathbf{c}_{j_2}, \ldots, \mathbf{c}_{j_k}\}$ denote the columns of R containing leading 1's. If we transpose the matrix, the same arguments as before show that $\{\mathbf{c}_{j_1}, \mathbf{c}_{j_2}, \ldots, \mathbf{c}_{j_k}\}$ is an independent set.

Let U denote the subspace of all columns in $\mathbb{R}^m$ in which the last $m-k$ entries are zero. Then $\dim U = k$ (it is just $\mathbb{R}^k$ with extra zeros). Since S contains the right number of independent vectors, then it is a basis by Theorem 13.23. This completes the proof. $\square$

EXAMPLE 14.8 (a) Find a subset of the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ -5 \\ 7 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -7 \\ -1 \end{pmatrix}, \qquad \mathbf{v}_3 = \begin{pmatrix} 1 \\ -4 \\ 2 \end{pmatrix}$$

that forms a basis for the space spanned by these vectors.

(b) Express each vector not in the basis as a linear combination of the basis vectors.

Solution (a) We begin by constructing a matrix that has $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ as its column vectors. The first part of our problem can be solved by finding a basis for the column space of this matrix. Reduce the matrix to a reduced row-echelon form by elementary row operations:

$$\begin{pmatrix} 2 & 1 & 1 \\ -5 & -7 & -4 \\ 7 & -1 & 2 \end{pmatrix} \to \begin{pmatrix} 2 & 1 & 1 \\ 3 & -3 & 0 \\ 3 & -3 & 0 \end{pmatrix} \to \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & 3 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Denote the column vectors of the resulting matrix by $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$. The pivots occur in columns 1 and 3, so, $\mathbf{w}_1$, $\mathbf{w}_3$ is a basis for the column space, and consequently, $\{\mathbf{v}_1, \mathbf{v}_3\}$ is a basis for the column space.

(b) Express each column vector that does not contain a pivot as a linear combination of column vectors that do contain pivots. This yields a set of **dependency equations**: $\mathbf{w}_2 = 3\mathbf{w}_3 - \mathbf{w}_1$. The corresponding equations for the column vectors of A express the vectors that are not in the basis as linear combinations of the basis vectors: $\mathbf{v}_2 = 3\mathbf{v}_3 - \mathbf{v}_1$. $\square$

Matrix Rank

THEOREM 14.9 (Rank Theorem) For any matrix A, the matrix rank is well-defined and is equal to the row rank and the column rank:

$$\operatorname{rank} A = \dim(\mathbf{col}\, A) = \dim(\mathbf{row}\, A).$$

Proof follows form Lemmas 14.4 and 14.7. □

EXAMPLE 14.10 Compute the rank of $A = \begin{pmatrix} 1 & 2 & 2 & -1 \\ 3 & 6 & 5 & 0 \\ 1 & 2 & 1 & 2 \end{pmatrix}$ and find bases for **row** A and **col** A.

Solution. The reduction of A to row-echelon form is as follows:

$$\begin{pmatrix} 1 & 2 & 2 & -1 \\ 3 & 6 & 5 & 0 \\ 1 & 2 & 1 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 2 & -1 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & -1 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 2 & -1 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence rank $A = 2$, and $\{[1 \quad 2 \quad 2 \quad -1], [0 \quad 0 \quad 1 \quad -3]\}$ is a basis of **row** A by Lemma 14.4. Since the leading 1's are in columns 1 and 3 of the row-echelon matrix, Theorem 14.5 shows that columns 1 and 3 of A form a basis of **col** A. □

Theorem 14.9 has several important consequences. The first, Corollary 14.11 below, follows because the rows of A are independent (respectively span **row** A) if and only if their transposes are independent (respectively span **col** A).

COROLLARY 14.11 If A is any matrix, then rank $A = \text{rank}(A^T)$.

What is the maximum possible rank of an $m \times n$ matrix A that is not square? The answer is given by the following Corollary.

COROLLARY 14.12 If A is an $m \times n$ matrix, then rank $A \leqslant m$ and rank $A \leqslant n$, that is,

$$\text{rank}(A) \leqslant \min(m, n).$$

Proof. Since the row vectors of A lie in $\mathbb{R}^n$, i.e. **row** $A \subset \mathbb{R}^n$, the row space of A is at most n-dimensional:

$$\dim(\mathbf{row}\, A) \leqslant \dim(\mathbb{R}^n) = n.$$

Since the column vectors lie in $\mathbb{R}^m$, i.e. **col** $A \subset \mathbb{R}^m$, the column space is at most m-dimensional:

$$\dim(\mathbf{col}\, A) \leqslant \dim(\mathbb{R}^m) = m.$$

As the rank of A is the common dimension of its row and column space, it follows that the rank is at most the smaller of m and n. □

The notion of rank of a matrix has a useful application to equations.

THEOREM 14.13 Suppose a system of m equations in n variables is consistent, and that the rank of the augmented matrix is r.

1. The set of solutions involves exactly $n - r$ parameters.
2. If $r < n$, the system has infinitely many solutions.
3. If $r = n$, the system has a unique solution.

Proof. The fact that the rank of the augmented matrix is r means there are exactly r leading variables, and hence exactly $n - r$ nonleading variables. These nonleading variables are all assigned as parameters in the Gaussian algorithm, so the set of solutions involves exactly $n-r$ parameters. Hence if $r<n$, there is at least one parameter, and so infinitely many solutions. If $r=n$, there are no parameters and so a unique solution. □

Size of Largest Non-vanishing Minor

DEFINITION 14.14 A **non-vanishing p-minor** is a submatrix of order p (i.e., $p\times p$ submatrix) of A with non-zero determinant. A **basis minor** is any non-vanishing minor of a largest order in A. A **determinantal rank** of A is an order of a basis minor.

THEOREM 14.15 (Basis Minor Theorem)

1. The rank of A is equal to the determinantal rank.
2. Rows of A passing through a basis minor form a basis in the **row** A.
3. Columns of A passing through a basis minor form a basis in the **col** A.

Proof. A non-vanishing p-minor shows that the rows and columns of that submatrix are linearly independent, and thus those rows and columns of the full matrix are linearly independent (in the full matrix), so the row and column rank are at least as large as the determinantal rank; however, the converse is less straightforward.

Show that the submatrix formed by basis rows and basis columns of A is non-vanishing. Suppose rank $A=r$ and consider a basis in the row space of A, containing r vectors. By relabelling, if necessary, the rows, we can assume that these are the first r rows. Subtract a linear combination of the basis rows from each row not in the basis so that it becomes a zero row. As elementary row operations do not alter dependence relationships among the column vectors, the reduced matrix R will have the same basis columns as A. Since the column vectors can be considered as vectors in R^r (with extra $m-r$ zeros), we see that the submatrix formed by basis columns is invertible (it follows from theory of determinants). □

This theorem does not give an efficient way of computing the rank, but it is useful theoretically: a single non-zero minor witnesses a lower bound (namely its order) for the rank of the matrix, which can be useful (for example) to prove that certain operations do not lower the rank of a matrix.

The next corollary requires a preliminary lemma.

LEMMA 14.16 Let A, B, and C be matrices of sizes $m\times n$, $n\times k$, and $l\times m$ respectively.

1. $\mathbf{col}(AB)\subset \mathbf{col}\,A$, with equality if B is (square and) invertible.
2. $\mathbf{row}(CA)\subset \mathbf{row}\,A$, with equality if C is (square and) invertible.

Proof. For (1), write $B=[\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_k]$ where $\mathbf{v}_j$ is column j of B. Then we have $AV=[A\mathbf{v}_1\ A\mathbf{v}_2\ \ldots\ A\mathbf{v}_k]$, and each $A\mathbf{v}_j$ is in $\mathbf{col}\,A$. It follows that $\mathbf{col}(AV)\subset\mathbf{col}\,A$. If B is invertible, we obtain $\mathbf{col}\,A=\mathbf{col}[(AV)V^{-1}]\subset\mathbf{col}(AV)$ in the same way. This proves (1).

As to (2), we have $\mathbf{col}[(CA)^T]=\mathbf{col}(A^TU^T)\subset\mathbf{col}(A^T)$ by (1), from which $\mathbf{row}(CA)\subset \mathbf{row}\,A$. If C is invertible, this is equality as in the proof of (1). □

COROLLARY 14.17 If A is $m \times n$ and B is $n \times k$, then $\operatorname{rank} AB \leqslant \operatorname{rank} A$ and $\operatorname{rank} AB \leqslant \operatorname{rank} B$.

Proof. By Lemma 14.15, $\mathbf{col}(AB) \subset \mathbf{col}\, A$ and $\mathbf{row}(BA) \subset \mathbf{row}\, A$, so Rank Theorem applies. $\square$

Exercises

1. Find the ranks of the following matrices:

$$\text{(a)} \begin{pmatrix} 0 & 0 & 0 & 7 \\ 0 & 0 & 6 & -1 \\ 0 & 13 & -18 & 3 \\ -4 & -5 & 4 & 0 \end{pmatrix}, \qquad \text{(b)} \begin{pmatrix} 8 & -11 & -7 & 2 \\ 0 & 3 & 8 & -1 \\ 0 & 15 & 40 & -5 \\ 0 & -6 & -16 & 2 \end{pmatrix}$$

Answer. (a) 2, (b) 2.

2. Find the ranks of the following matrices:

$$\text{(a)} \begin{pmatrix} 9 & 4 & 34 & 47 \\ 2 & 1 & 8 & 11 \\ -13 & -5 & -46 & -64 \\ 11 & 6 & 46 & 63 \end{pmatrix}, \qquad \text{(b)} \begin{pmatrix} -6 & -1 & 2 & 3 & 1 \\ -5 & 0 & 2 & 4 & 3 \\ -12 & -2 & 5 & 6 & -1 \\ 6 & 1 & -2 & -2 & 2 \end{pmatrix}$$

Answer. (a) 2, (b) 4.

3. Determine the rank of the following matrix for various values of the parameter:

$$\text{(a)} \begin{pmatrix} 1 & 1 & 1 \\ 1 & a & a \\ 1 & a^2 & a^2 \end{pmatrix}, \qquad \text{(b)} \begin{pmatrix} 1 & 1 & 1 \\ 1 & b & b^2 \\ 1 & b^2 & b \end{pmatrix}, \qquad \text{(c)} \begin{pmatrix} 1 & \alpha & -1 & 2 \\ 2 & -1 & \alpha & 5 \\ 1 & 10 & -6 & 1 \end{pmatrix}$$

Answer. (a) 1 if $a = 1$; 2 otherwise, (b) 1 if $b = 1$; 2 if $b = 0$ and $b = -2$; 3 otherwise, (c) 2 if $\alpha = 3$; 3 otherwise.

4. Find a subset of the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \\ -1 \\ -8 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 5 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 3 \\ -1 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix}, \quad \mathbf{v}_5 = \begin{pmatrix} 0 \\ -3 \\ 2 \\ 5 \end{pmatrix}$$

that forms a basis for the space spanned by these vectors. Express each vector that is not in the basis as a linear combination of the basis vectors.

Answer. $\{\mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is a basis, $\mathbf{v}_1 = \mathbf{v}_2 + \mathbf{v}_3 - 3\mathbf{v}_4$, $\mathbf{v}_5 = 2\mathbf{v}_4 - \mathbf{v}_2$.