The Column Space & Column Rank of a Matrix

E. L. Lady

Let $A = \begin{bmatrix} 0 & 1 & 2 & -3 & 4 \\ 0 & 2 & 4 & 0 & 20 \\ 0 & -1 & -2 & 5 & 0 \end{bmatrix}$.

The **column space** of A is the subspace of $\mathbb{R}^3$ spanned by the columns of A, in other words it consists of all linear combinations of the columns of A:

$$u \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + w \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + x \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} + y \begin{bmatrix} -3 \\ 0 \\ 5 \end{bmatrix} + z \begin{bmatrix} 4 \\ 20 \\ 0 \end{bmatrix}.$$

It can then be seen that a vector $\mathbf{b}$ with coordinates b_1, b_2 and b_3 belongs to the column space of A precisely when the system of equations

$$\begin{bmatrix} 0 & 1 & 2 & -3 & 4 \\ 0 & 2 & 4 & 0 & 20 \\ 0 & -1 & -2 & 5 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \mathbf{b}$$

has a solution.

Now if we take the augmented matrix for this linear system and reduce it to row echelon form then we get

$$\begin{bmatrix} 0 & 1 & 2 & -3 & 4 & | & b_1' \\ 0 & 0 & 0 & 1 & 2 & | & b_2' \\ 0 & 0 & 0 & 0 & 0 & | & b_3' \end{bmatrix}$$

(where b_1', b_2', and b_3' are new constants).

For instance, if we are trying to solve

$$\begin{bmatrix} 0 & 1 & 2 & -3 & 4 \\ 0 & 2 & 4 & 0 & 20 \\ 0 & -1 & -2 & 5 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ -6 \\ -7 \end{bmatrix}$$

then the augmented matrix reduces to the row echelon form

$$\left[\begin{array}{ccccc|c} 0 & 1 & 2 & -3 & 4 & 3 \\ 0 & 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

and so the solutions have the form

$$\begin{bmatrix} u \\ w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 0 \\ -2 \\ 0 \end{bmatrix} + r \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -10 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

where r, s, and t are arbitrary constants.

> **Important:** For instance, we can get one solution (a particular solution) by setting $r = s = t = 0$. In general, **if a solution exists at all, then there is a solution in which all variables are zero which correspond to columns which do NOT contain leading entries in the row echelon form.** (Be sure that you understand why this is true!)

By choosing the solution with $r = s = t = 0$ we get

$$\begin{bmatrix} 3 \\ -6 \\ -7 \end{bmatrix} = 0 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - 3 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} - 2 \begin{bmatrix} -3 \\ 0 \\ 5 \end{bmatrix} + 0 \begin{bmatrix} 4 \\ 20 \\ 0 \end{bmatrix}.$$

Applying this in general, we can see that **if** $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ is in the column space of A, then it is possible to write

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = 0 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + w \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} + y \begin{bmatrix} -3 \\ 0 \\ 5 \end{bmatrix} + 0 \begin{bmatrix} 4 \\ 20 \\ 0 \end{bmatrix}$$

for the right choice of w and y, so that $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ can be written as a linear combination of the two columns $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 0 \\ 5 \end{bmatrix}$.

Restated, this says that

> The column space of A is spanned by the two columns
> $$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \text{ and } \begin{bmatrix} -3 \\ 0 \\ 5 \end{bmatrix}.$$

Since it can be easily (!) seen that these two vectors are linearly independent, it follows that they are a **basis** for the column space of A. Therefore the **column rank** of A, which by definition is the dimension of the column space, is 2.

As a general principle, let A by any $m \times n$ matrix and let $A_1, \ldots, A_n$ be the columns of A, so that $A = [\,A_1 \quad A_2 \quad \ldots \quad A_n\,]$. Then these n columns A_i are vectors in $\mathbb{R}^m$ and the column space of A consists of all vectors $\mathbf{b} \in \mathbb{R}^m$ which can be written in the form

$$\mathbf{b} = x_1 A_1 + \cdots + x_n A_n = A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Now the procedure for solving systems of linear equations by elimination shows that if it is possible to solve the system

$$A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{b}$$

at all, then there is one solution where **only** those variables $x_{j_1}, \ldots, x_{j_r}$ corresponding to columns containing the leading entry for some row in the row echelon form of A need to be non-zero. (So here r will be the number of leading entries in the row echelon form, which is the same as the number of non-zero rows in the row echelon form.) This says that if $\mathbf{b}$ is in the column space of A then there exist values $x_{j_1}, \ldots, x_{j_r}$ such that $\mathbf{b} = x_{j_1} A_{j_1} + \cdots + x_{j_r} A_{j_r}$.

Restated, this says that the column space of A is spanned by the columns $A_{j_1}, \ldots, A_{j_r}$, where $j_1, \ldots, j_r$ are the numbers for the columns which contain the leading entry for some row in the row echelon form of A. (Thus r is the number of non-zero rows in the row echelon form of A.)

Furthermore, these columns $A_{j_1}, \ldots, A_{j_r}$ will always be linearly independent. In fact, we can see that $x_{j_1} A_{j_1} + \cdots + x_{j_r} A_{j_r} = \mathbf{0}$ is only possible when $x_{j_1} = \cdots = x_{j_r} = 0$ by noticing that the row echelon form of $[A_{j_1} \ldots A_{j_r}]$ is an $m \times r$ matrix whose row echelon form has a leading entry in each column. (Why? Hint: How does the row echlon form for $[A_{j_1} \ldots A_{j_r}]$ compare to the row echelon form of $A = [A_1 \ldots A_n]$?)