The column space and null space of a matrix

Sean Fitzpatrick

2015-02-25

In our course, we've discussed the general scenario of the null space and range of a linear transformation between vector spaces. Given a linear transformation $T : V \to W$, our main results are that $\operatorname{null} T$ is a subspace of V, and $\operatorname{range} T$ is a subspace of W, and if V is finite-dimensional, then

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

As mentioned in class, a special case of linear transformations in general is that of a matrix transformation. If we denote

$$V = \mathbb{R}^{n,1} = \left\{ X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} : x_1, x_2, \ldots, x_n \in \mathbb{R} \right\}$$

and

$$W = \mathbb{R}^{m,1} = \left\{ Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ v_m \end{bmatrix} : y_1, y_2, \ldots, y_m \in \mathbb{R} \right\},$$

and if A is any $m \times n$ matrix, then we have a corresponding linear transformation $T_A : \mathbb{R}^{n,1} \to \mathbb{R}^{m,1}$ given by

$$T_A(X) = AX = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Notice that $\mathcal{M}(T_A) = A$: the matrix of the linear transforation T_A, as defined in the text, is equal to the original matrix A. This implies that for finite-dimensional vector spaces, the properties of general linear transformations are closely related to those of matrix transformations. However, the study of matrix transformations in particular, as opposed to linear transformations in general, often falls through the cracks: it belongs in the study of vectors in $\mathbb{R}^n$ covered in a course like Math 1410, but is usually left out due to time constraints, and doesn't quite make it into a course like Math 3410 that begins with the study of general vector spaces.

Given an $m \times n$ matrix A, one defines the *null space* of A to be the subspace of $\mathbb{R}^{n,1}$ defined by

$$\operatorname{null} A = \{X \in \mathbb{R}^{n,1} : AX = 0\},$$

and the *column space* of A to be the subspace $\operatorname{col} A$ of $\mathbb{R}^{m,1}$ spanned by the columns of A. However, if we follow the usual rules of matrix multiplication, we can see that

$$AX = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Thus, $Y \in \mathbb{R}^{m,1}$ belongs to the column space of A if and only if $Y = AX$ for some vector $X \in \mathbb{R}^{n,1}$. In other words,

$$\operatorname{col} A = \operatorname{range} T_A.$$

In a course focusing on matrices and vectors in $\mathbb{R}^n$, one is usually concerned with the rank of a matrix, since this gives us information about the general solutions to systems of linear equations. One often defines in addition row A, the *row space* of A, to be the subspace of $\mathbb{R}^{1,m}$ spanned by the rows of A.[1]

1. The **rank** of A, given by the number of leading ones in the row-echelon form of A.
2. The **column rank** of A, given by the dimension of $\operatorname{col} A$.

[1]Note that $\mathbb{R}^{m,1}$ and $\mathbb{R}^{1,m}$ are isomorphic: the transpose of a row is a column, and vice versa. In fact there is an even closer relationship here: $\mathbb{R}^{1,m}$ is the *dual* of $\mathbb{R}^{m,1}$. If you have the third edition of the text, you can find details in section 3.F. In particular, you'll find there the result that given a linear transformation $T : V \to W$ between finite-dimensional vector spaces V and W, with dual spaces V' and W', there is a dual map $T' : W' \to V'$, and the matrix of T' is simply the transpose of the matrix of T.

3. The **row rank** of A, given by the dimension of row A.

One then proves that the rank, column rank, and row rank of A are all equal. The proof is essentially the Gaussian elimination algorithm: the leading ones in the row-echelon form of A each land in exactly one row and one column.

The next few results are borrowed from the text "Linear Algebra with Applications", 5^{th} ed., by W. Keith Nicholson.

Theorem 1. *Let A, U, and V be matrices of size $m\times n$, $p\times m$, and $n\times q$, respectively. Then*

1. $\operatorname{col}(AV) \subseteq \operatorname{col} A$, *with equality if V is square and invertible.*
2. $\operatorname{row}(UA) \subseteq \operatorname{row} A$, *with equality if U is square and invertible.*

Proof. We will prove (1); the proof of (2) follows from (1) by taking transposes. Note that if X_j denotes the j^{th} column of V, then $AX_j \in \operatorname{col} A$. But AX_j is just the j^{th} column of AV. Since $\operatorname{col} A$ contains a spanning set for $\operatorname{col}(AV)$, it follows that $\operatorname{col}(AV) \subseteq \operatorname{col} A$. If V is invertible, then

$$\operatorname{col}(AV) \supseteq \operatorname{col}((AV)V^{-1}) = \operatorname{col} A,$$

so $\operatorname{col}(AV) = \operatorname{col} A$. $\square$

Note that this result is almost trivial from the point of view of linear transformations, using basic properties of functions from Math 2000: we have $\operatorname{col} A = \operatorname{range} T_A$, and $\operatorname{col}(AV) = \operatorname{range} T_{AV} = \operatorname{range}(T_A \circ T_V)$. Since $\operatorname{range} T_V \subseteq \operatorname{dom} T_A$, it follows that $\operatorname{range} T_A \circ T_V \subseteq \operatorname{range} T_A$.

We now want to consider the above from the point of view of the row-echelon form of a matrix A. Recall that since A can be carried to its row-echelon form R (which is unique) via elementary row operations, there exist elementary matrices $E_1, \ldots, E_K$ such that

$$R = (E_k \cdots E_1)A = UA,$$

where $U = E_k \cdots E_1$ is a product of elementary matrices, and therefore invertible. It immediately follows from the above theorem that $\operatorname{row} R = \operatorname{row} A$. In addition, we have the following:

Lemma 1. *Let R denote an $m \times n$ matrix in row-echelon form. Then:*

1. *The non-zero rows of R are a basis for* $\operatorname{row} R$.
2. *The columns of R containing the leading ones are a basis cor* $\operatorname{col} R$.

Proof.

1. By definition, $\operatorname{row} R = \operatorname{span}\{R_1, \ldots, R_r\}$, where $r = \operatorname{rank} R$ and the R_j are the nonzero rows of R. From the definition of row-echelon form, we can assume that the rows are ordered such that the leading one of R_i is to the left of the leading one of R_j if $i < j$. It is then easy to see that the rows are linearly independent: if $a_1 R_1 + \cdots + a_m R_m = 0$, then $a_1 = 0$, since the leading one of R_1 is to the left of the remaining leading ones, and then $a_2 = 0$, and so on.

2. The columns containing leading ones are linearly independent, since each contains a leading one in a different row, and all the entries below a leading one are zero, so $\dim \operatorname{col} R \geq r$. Note from the proof of (1) that only the first r rows of R contain nonzero entries, so $\operatorname{col} R$ is contained in the subspace of $\mathbb{R}^m$ consisting of columns with the last $m - r$ entries equal to zero. This shows that $\dim \operatorname{col} R \leq r$, and that the columns containing the leading ones must therefore form a basis.

□

Note that in the course of the above proof, we've shown that for a matrix R in row-echelon form, $\dim \operatorname{row} R = \dim \operatorname{col} R = \operatorname{rank} R$. It remains to show that this still holds true for any matrix.

Theorem 2. *Let A denote any $m \times n$ matrix. Then*

$$\dim \operatorname{row} A = \dim \operatorname{col} A.$$

Moreover, suppose that A can be carried to a matrix R in row-echelon form via elementary row operations. If $r = \operatorname{rank} A$ is the number of nonzero rows in R, then

1. *The r nonzero rows of R form a basis for* $\operatorname{row} A$.

2. *If the leading ones of R appear in columns $j_1, \ldots, j_r$ of R, then the corresponding columns $j_1, \ldots, j_r$ of A are a basis for* $\operatorname{col} A$.

Remark: It is the second point in the above theorem that is the most relevant to us. (Knowing the row space is not so important.) If $A = \mathcal{M}(T)$ denotes the matrix of a linear transformation T, then knowing $\operatorname{col} A$ will let us determine $\operatorname{range} T$. Be careful with this result though: it is **not** the leading columns of R that form a basis for $\operatorname{col} A$: the leading ones in R just tell us which columns *of the original matrix A* will form a basis for the column space of A.

Proof of Theorem 2. We know that $R = UA$ for some invertible matrix U (since U will be a product of elementary matrices). It follows that $\operatorname{row} A = \operatorname{row} R$ from

Theorem 1, and (1) follows from the lemma above. Now, let $C_1, \dots, C_n$ denote the columns of A, and write $A = \begin{bmatrix} C_1 & \cdots & C_n \end{bmatrix}$ in block form. Then

$$R = UA = \begin{bmatrix} UC_1 & \cdots & UC_n \end{bmatrix}.$$

If $j_1, \dots, j_r$ are the columns of R containing leading ones, then by our lemma, the columns $UC_{j_1}, \dots, UC_{j_r}$ of R form a basis for $\operatorname{col} R$. Since these columns are linearly independent and U is invertible, it follows that the columns $C_{j_1}, \dots, C_{j_r}$ of A are linearly independent.[2] Finally, for any column C_j of A, we know that UC_j is a linear combination of $UC_{j_1}, \dots, UC_{j_r}$, since these vectors form a basis for $\operatorname{col} R$, and since U is invertible, it follows that C_j is a linear combination of $C_{j_1}, \dots, C_{j_r}$, and we've proved (2).

As a consequence of (1) and (2), we have that $\dim \operatorname{col} A = \dim \operatorname{row} A = r = \operatorname{rank} A$. □

Corollary 1. *For any matrix A,* $\operatorname{rank} A = \operatorname{rank} A^T$.

Corollary 2. *If A is an $m \times n$ matrix, then* $\operatorname{rank} A \leq \min\{m, n\}$.

Corollary 3. *An $n \times n$ matrix A is invertible if and only if* $\operatorname{rank} A = n$.

Another consequence of the above is the so-called "Rank-Nullity Theorem" for a matrix. If we define the nullity of A to be the dimension of $\operatorname{null} A$, then this theorem states that for any $m \times n$ matrix A,

$$\operatorname{nullity} A + \operatorname{rank} A = n.$$

If we view A as the coefficient matrix for a system of m equations in n variables, this is the familiar result that the number of parameters in the general solution to the system (given by the nullity of A) is equal to the number of variables (n) minus the rank. However, if we view A as the matrix of a linear transformation $T : V \to W$, then we see that $\operatorname{rank} A = \dim \operatorname{col} A = \dim \operatorname{range} A$, and $\operatorname{nullity} A = \dim \operatorname{null} A = \dim \operatorname{null} T$, then this is again the "Fundamental Theorem of Linear Maps", that

$$\dim \operatorname{null} A + \dim \operatorname{range} A = \dim V.$$

[2] This is a familar result in different language. If we consider $C_{j_1}, \dots, C_{j_r}$ as vectors in $\mathbb{R}^{m,1}$, and multiplication by U as an invertible (and thus, injective) linear transformation $T_U : \mathbb{R}^{m,1} \to \mathbb{R}^{m,1}$, we know that the vectors $UC_{j_1}, \dots, UC_{j_r}$ are linearly independent if and only if the vectors $C_{j_1}, \dots, C_{j_r}$ are linearly independent.