

Classroom



In this section of *Resonance*, we invite readers to pose questions likely to be raised in a classroom situation. We may suggest strategies for dealing with them, or invite responses, or both. “Classroom” is equally a forum for raising broader issues and sharing personal experiences and viewpoints on matters related to teaching and learning science.

On the Rank of a Matrix*

Two important results in linear algebra are the ‘rank-nullity theorem’ and the equality of the row and column ranks of a matrix. In this note, we will give a simple proof of the latter, using the former. As a by-product, we also prove the Fredholm alternative, which characterizes the range of the linear operator associated with a matrix.

In the Euclidean space, \mathbb{R}^n , we will denote the usual inner-product by (\cdot, \cdot) and the corresponding euclidean norm by $|\cdot|$. Given an $m \times n$ matrix A , we will denote its transpose by A^T . In particular, given a column vector $\mathbf{x} \in \mathbb{R}^n$, the symbol \mathbf{x}^T will denote the corresponding row vector and *vice versa*. In general, we will treat vectors in \mathbb{R}^n as column vectors. Thus, if $\mathbf{x} = (x_1, \dots, x_n)^T$ and $\mathbf{y} = (y_1, \dots, y_n)^T$ are vectors in \mathbb{R}^n , we have

$$(\mathbf{x}, \mathbf{y}) = \mathbf{y}^T \mathbf{x} = \sum_{i=1}^n x_i y_i, \text{ and } |\mathbf{x}|^2 = \sum_{i=1}^n |x_i|^2.$$

If A is an $m \times n$ matrix, then it can be considered as a linear transformation mapping \mathbb{R}^n into \mathbb{R}^m , and we will denote it by the

S Kesavan
Adjunct Faculty
Department of Mathematics
Indian Institute of
Technology-Madras
IIT P.O., Chennai 600 036
India.
Email: kesh@imsc.res.in

Two important results in linear algebra are the ‘rank-nullity theorem’ and the equality of the row and column ranks of a matrix.

Keywords

Rank of a matrix, range of a linear operator, Fredholm alternative, rank-nullity theorem.

*Vol.26, No.4, DOI: <https://doi.org/10.1007/s12045-021-1156-6>

symbol \mathbf{A} . The rank-nullity theorem states that

$$\dim(R(\mathbf{A})) + \dim(N(\mathbf{A})) = n, \quad (1)$$

where \dim denotes the dimension of a vector space, $R(\mathbf{A})$ denotes the range (*i.e.* the image) of the operator \mathbf{A} (which is a subspace of \mathbb{R}^m), and $N(\mathbf{A})$ denotes the null-space of \mathbf{A} , *i.e.* the subspace of \mathbb{R}^n of vectors \mathbf{x} such that $\mathbf{A}\mathbf{x} = \mathbf{0}$.

In particular, if $n = m$, then the above result implies that \mathbf{A} is injective if, and only if, it is surjective, in which case it is invertible.

Let W be a subspace of \mathbb{R}^n . Given $\mathbf{x} \in \mathbb{R}^n$, there exists a unique vector, denoted $\mathbf{P}\mathbf{x} \in W$ such that

$$(\mathbf{P}\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{y}) \text{ for all } \mathbf{y} \in W.$$

To see this, first notice that, by linearity, it is enough to check the above relation for just the basis vectors of W . Thus, if $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ is a basis for W , it is enough to find $\mathbf{P}\mathbf{x}$ such that

$$(\mathbf{P}\mathbf{x}, \mathbf{w}_i) = (\mathbf{x}, \mathbf{w}_i) \text{ for all } 1 \leq i \leq k.$$

We write $\mathbf{P}\mathbf{x} = \sum_{j=1}^k \alpha_j \mathbf{w}_j$, so that we get k linear equations in k unknowns,

$$\sum_{j=1}^k (\mathbf{w}_j, \mathbf{w}_i) \alpha_j = (\mathbf{x}, \mathbf{w}_i) \text{ for all } 1 \leq i \leq k.$$

The $k \times k$ matrix associated with this system, $G = (g_{ij})$, where $g_{ij} = (\mathbf{w}_j, \mathbf{w}_i)$, satisfies, for every $\mathbf{z} \in \mathbb{R}^k$,

$$\mathbf{z}^T G \mathbf{z} = \sum_{i,j=1}^k g_{ij} z_i z_j = \left| \sum_{i=1}^k z_i \mathbf{w}_i \right|^2,$$

which is strictly positive if $\mathbf{z} \neq \mathbf{0}$. In particular, this matrix defines an injective linear transformation and hence, by our earlier observation, it is also invertible. Thus, the system of equations above has a unique solution, which establishes the existence of $\mathbf{P}\mathbf{x}$.

The mapping $\mathbf{x} \mapsto \mathbf{P}\mathbf{x}$ defines a linear transformation of \mathbb{R}^n into itself and, clearly, $R(\mathbf{P}) = W$. Then

$$N(\mathbf{P}) = \{\mathbf{z} \in \mathbb{R}^n \mid (\mathbf{z}, \mathbf{y}) = 0 \text{ for all } \mathbf{y} \in W\},$$



which we denote as W^\perp . The subspace W^\perp is called the orthogonal complement of the subspace W . By the rank-nullity theorem,

$$\dim(W) + \dim(W^\perp) = n. \quad (2)$$

Let A be an $m \times n$ matrix with real entries. The row (respectively, column) rank of the matrix A is the maximum number of linearly independent rows (respectively, columns) in A . It is then clear that the column rank is just the dimension of the range of the transformation \mathbf{A} and the row rank is the dimension of the range of the transformation \mathbf{A}^T . As mentioned earlier, an important theorem in linear algebra is that the row and column ranks of a matrix are, in fact, equal, and the common value is called the rank of the matrix. We will now prove this fact.

The row and column ranks of a matrix are, in fact, equal, and the common value is called the rank of the matrix.

Let $\mathbf{b} \in R(\mathbf{A})$. Then, there exists $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}$. Let $\mathbf{y} \in N(\mathbf{A}^T)$, i.e. $\mathbf{A}^T\mathbf{y} = \mathbf{0}$. Then

$$(\mathbf{y}, \mathbf{b}) = \mathbf{y}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{A}^T\mathbf{y} = 0.$$

It follows from this that $R(\mathbf{A}) \subset N(\mathbf{A}^T)^\perp$. Consequently, (using (2) and (1), we get

$$\dim(R(\mathbf{A})) \leq \dim(N(\mathbf{A}^T)^\perp) = m - \dim(N(\mathbf{A}^T)) = \dim(R(\mathbf{A}^T)). \quad (3)$$

Applying this to \mathbf{A}^T , we get that

$$\dim(R(\mathbf{A}^T)) \leq \dim(R(\mathbf{A})),$$

from which we deduce that

$$\dim(R(\mathbf{A})) = \dim(R(\mathbf{A}^T)),$$

which, by our earlier observation, is just the fact that the column rank of A is equal to its row rank.

Since we now have equality throughout in (3), it follows that

$$R(\mathbf{A}) = N(\mathbf{A}^T)^\perp,$$

which characterises the range of the transformation \mathbf{A} . This result is often referred to in the literature as the *Fredholm alternative*.



CLASSROOM

The case of complex matrices follows the same argument *mutatis mutandis*. In this case, the above proof shows that the column rank of A is the column rank of its conjugate transpose A^* . But then it is immediate to see that the transpose A^T and the conjugate transpose A^* have the same column ranks and this completes the proof in the complex case.

