MA 110:
Lecture 05

Saurav Bhaumik
Department of Mathematics
IIT Bombay

Spring 2025

Recall:

Proposition

An $n \times n$ matrix is invertible if and only if it can be transformed to the $n \times n$ identity matrix by EROs.

Proof. Suppose $\mathbf{A} \in \mathbb{R}^{n \times n}$ is invertible. Using EROs, transform $\mathbf{A}$ to a matrix $\mathbf{A}' \in \mathbb{R}^{n \times n}$ such that $\mathbf{A}'$ is in a RCF. Since $\mathbf{A}$ is invertible, the linear system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution. Hence $\mathbf{A}'$ has n nonzero rows, and so each of the n columns of $\mathbf{A}'$ is pivotal. Also, the number of rows of $\mathbf{A}$ is equal to the number of its columns, that is, $m = n$. Therefore $\mathbf{A}' = \mathbf{I}$.

Conversely, suppose $\mathbf{A} \in \mathbb{R}^{n \times n}$ can be transformed to the $n \times n$ identity matrix $\mathbf{I}$ by EROs. Since $\mathbf{Ix} = \mathbf{0} \implies \mathbf{x} = \mathbf{0}$ for $\mathbf{x} \in \mathbb{R}^{n \times 1}$, we see that the linear system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution. Hence $\mathbf{A}$ is invertible. $\square$

Remark.

Suppose an $n \times n$ square matrix $\mathbf{A}$ is invertible. In order to solve the linear system $\mathbf{Ax} = \mathbf{b}$ for a given $\mathbf{b} \in \mathbb{R}^{n\times 1}$, we may transform the augmented matrix $[\mathbf{A}|\mathbf{b}]$ to $[\mathbf{I}\,|\,\mathbf{c}]$ by EROs. Now $\mathbf{Ax} = \mathbf{b} \iff \mathbf{Ix} = \mathbf{c}$ for $\mathbf{x} \in \mathbb{R}^{n\times 1}$. Hence $\mathbf{Ac} = \mathbf{b}$. Thus $\mathbf{c}$ is the unique solution of $\mathbf{Ax} = \mathbf{b}$. This observation is the basis of an important method to find the inverse of a square matrix.

Gauss-Jordan Method for Finding the Inverse of a Matrix

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be an invertible matrix. Consider the basic column vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^{n\times 1}$. Then $\begin{bmatrix}\mathbf{e}_1 & \cdots & \mathbf{e}_n\end{bmatrix} = \mathbf{I}$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be the unique elements of $\mathbb{R}^{n\times 1}$ be such that $\mathbf{Ax}_1 = \mathbf{e}_1, \ldots, \mathbf{Ax}_n = \mathbf{e}_n$, and define $\mathbf{X} := \begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix}$. Then

$$\mathbf{AX} = \mathbf{A}\begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix} = \begin{bmatrix}\mathbf{Ax}_1 & \cdots & \mathbf{Ax}_n\end{bmatrix} = \begin{bmatrix}\mathbf{e}_1 & \cdots & \mathbf{e}_n\end{bmatrix} = \mathbf{I}.$$

By an earlier result, it follows that $\mathbf{X} = \mathbf{A}^{-1}$.

Hence to find $\mathbf{A}^{-1}$, we may solve the n linear systems $\mathbf{A}\mathbf{x}_1 = \mathbf{e}_1, \ldots, \mathbf{A}\mathbf{x}_n = \mathbf{e}_n$ simultaneously by considering the $n \times 2n$ augmented matrix

$$[\mathbf{A}|\mathbf{e}_1 \cdots \mathbf{e}_n] = [\mathbf{A} \,|\, \mathbf{I}]$$

and transform $\mathbf{A}$ to its RCF, namely to $\mathbf{I}$, by EROs. Thus if $[\mathbf{A} \,|\, \mathbf{I}]$ is transformed to $[\mathbf{I} \,|\, \mathbf{X}]$, then $\mathbf{X}$ is the inverse of $\mathbf{A}$.

Remark To carry out the above process, we need not know beforehand that the matrix $\mathbf{A}$ is invertible. This follows by noting that $\mathbf{A}$ can be transformed to the identity matrix by EROs if and only if $\mathbf{A}$ is invertible. Hence the process itself reveals whether $\mathbf{A}$ is invertible or not.

Example

Let

$$\mathbf{A} := \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}.$$

We use EROs to transform $[\mathbf{A} \mid \mathbf{I}]$ to $[\mathbf{I} \mid \mathbf{X}]$, where $\mathbf{X} \in \mathbb{R}^{3\times 3}$.

$$\left[\begin{array}{ccc|ccc} -1 & 1 & 2 & 1 & 0 & 0 \\ 3 & -1 & 1 & 0 & 1 & 0 \\ -1 & 3 & 4 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 2 & 2 & -1 & 0 & 1 \end{array}\right] \to$$

$$\left[\begin{array}{ccc|ccc} \boxed{-1} & 1 & 2 & 1 & 0 & 0 \\ 0 & \boxed{2} & 7 & 3 & 1 & 0 \\ 0 & 0 & \boxed{-5} & -4 & -1 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & -1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 3.5 & 1.5 & 0.5 & 0 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0.6 & 0.4 & -0.4 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -0.7 & 0.2 & 0.3 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] = [\mathbf{I} \mid \mathbf{X}].$$

Thus **A** is invertible and

$$\mathbf{A}^{-1} = \mathbf{X} = \frac{1}{10}\begin{bmatrix} -7 & 2 & 3 \\ -13 & -2 & 7 \\ 8 & 2 & -2 \end{bmatrix}.$$

Linear Dependence

Let $n \in \mathbb{N}$. We shall work entirely with row vectors in $\mathbb{R}^{1\times n}$ (of length n), or entirely with column vectors in $\mathbb{R}^{n\times 1}$(of length n), both of which will be referred to as 'vectors'.

We have already considered a linear combination

$$\alpha_1 \mathbf{a}_1 + \cdots + \alpha_m \mathbf{a}_m$$

of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$, where $\alpha_1, \ldots, \alpha_m$ are scalars.

A set S of vectors is called **linearly dependent** if there is $m \in \mathbb{N}$, there are (distinct) vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$ in S and there are scalars $\alpha_1, \ldots, \alpha_m$, not all zero, such that

$$\alpha_1 \mathbf{a}_1 + \cdots + \alpha_m \mathbf{a}_m = \mathbf{0}.$$

It can be seen that S is linearly dependent $\iff$ either $\mathbf{0} \in S$ or a vector in S is a linear combination of other vectors in S.

Examples

(i) Let $S := \left\{ \begin{bmatrix} 1 & 2 \end{bmatrix}^{\mathsf{T}}, \begin{bmatrix} 2 & 1 \end{bmatrix}^{\mathsf{T}}, \begin{bmatrix} 1 & -1 \end{bmatrix}^{\mathsf{T}} \right\} \subset \mathbb{R}^{2\times 1}$. Then the set S is linearly dependent since

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \text{ Clearly, } \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

(ii) Let
$S := \left\{ \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}, \begin{bmatrix} 2 & 3 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & -3 & 3 \end{bmatrix} \right\} \subset \mathbb{R}^{1\times 3}$.
Then the set S is linearly dependent since
$\begin{bmatrix} 0 & -3 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} - 2\begin{bmatrix} 2 & 3 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 1 & 2 \end{bmatrix}$. Clearly,

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} - 2\begin{bmatrix} 2 & 3 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 1 & 2 \end{bmatrix} - \begin{bmatrix} 0 & -3 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}.$$

In (i) above, S is a set of 3 vectors in $\mathbb{R}^{2\times 1}$, and in (ii) above, S is a set of 4 vectors in $\mathbb{R}^{1\times 3}$. These examples illustrate an important phenomenon to which we now turn. First we prove the following crucial result.

Proposition

Let S be a set of s vectors, each of which is a linear combination of elements of a (fixed) set of r vectors. If $s > r$, then the set S is linearly dependent.

Proof. Let $S := \{\mathbf{x}_1, \ldots, \mathbf{x}_s\}$, and suppose each vector in S is a linear combination of elements of the set $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$ of r vectors and $s > r$. Then

$$\mathbf{x}_j = \sum_{k=1}^{r} a_{jk}\mathbf{y}_k \quad \text{for } j = 1, \ldots, s, \text{ where } a_{jk} \in \mathbb{R}.$$

Let $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{s\times r}$. Then $\mathbf{A}^\mathsf{T} \in \mathbb{R}^{r\times s}$. Since $r < s$, the linear system $\mathbf{A}^\mathsf{T}\mathbf{x} = \mathbf{0}$ has a nonzero solution, that is, there are $\alpha_1, \ldots, \alpha_s$, not all zero, such that

$$\mathbf{A}^\mathsf{T}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_s\end{bmatrix} = \begin{bmatrix}a_{11} & \cdots & a_{s1}\\ \vdots & \vdots & \vdots\\ a_{1r} & \cdots & a_{sr}\end{bmatrix}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_s\end{bmatrix} = \begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix} \in \mathbb{R}^{r\times 1},$$

that is, $\sum_{j=1}^{s} a_{jk}\alpha_j = 0$ for $k = 1, \ldots, r$. It follows that

$$\sum_{j=1}^{s} \alpha_j \mathbf{x}_j = \sum_{j=1}^{s} \alpha_j \left(\sum_{k=1}^{r} a_{jk} \mathbf{y}_k \right) = \sum_{k=1}^{r} \left(\sum_{j=1}^{s} a_{jk} \alpha_j \right) \mathbf{y}_k = \mathbf{0}.$$

Since not all $\alpha_1, \ldots, \alpha_n$ are zero, S is linearly dependent. □

Corollary

Let $n \in \mathbb{N}$ and S be a set of vectors of length n. If S has more than n elements, then S is linearly dependent.

Proof. If S is a set of column vectors of length n, then each element of S is a linear combination of the n column vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Similarly, if S is a set of row vectors of length n, then each element of S is a linear combination of the n row vectors $\mathbf{e}_1^{\mathsf{T}}, \ldots, \mathbf{e}_n^{\mathsf{T}}$. Hence the desired result follows from the crucial result we just proved. □

Linear Independence

A set S of vectors is called **linearly independent** if it is not linearly dependent, that is,

$$\alpha_1 \mathbf{a}_1 + \cdots + \alpha_m \mathbf{a}_m = \mathbf{0} \implies \alpha_1 = \cdots = \alpha_m = 0,$$

whenever $\mathbf{a}_1, \ldots, \mathbf{a}_m$ are (distinct) vectors in S and $\alpha_1, \ldots, \alpha_m$ are scalars. We may also say that the vectors in S are linearly independent.

Linearly independent sets are important because each one of them gives us data that we cannot obtain from any linear combination of the others. In this sense, each element of a linearly independent set is indispensable!

Examples

(i) The empty set is linearly independent vacuously.

(ii) Let S be the subset of $\mathbb{R}^{n\times 1}$ consisting of the basic column vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Then S is linearly independent. To see this, let $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $\alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n = \mathbf{0}$. Then the jth entry α_j of $\alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n$ is equal to 0 for $j = 1, \ldots, n$.

(iii) Let S denote the subset of $\mathbb{R}^{1\times 4}$ consisting of the vectors $\begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}, \begin{bmatrix}1 & 1 & 0 & 0\end{bmatrix}, \begin{bmatrix}1 & 1 & 1 & 0\end{bmatrix}$ and $\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix}$.

Then S is linearly independent. To see this, let $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$ be such that

$\alpha_1 \begin{bmatrix}1 & 0 & \cdots & 0\end{bmatrix} + \alpha_2 \begin{bmatrix}1 & 1 & 0 & 0\end{bmatrix} + \alpha_3 \begin{bmatrix}1 & 1 & 1 & 0\end{bmatrix} + \alpha_4 \begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix} = \begin{bmatrix}0 & 0 & 0 & 0\end{bmatrix}$.

Then $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 0$, $\alpha_2 + \alpha_3 + \alpha_4 = 0$, $\alpha_3 + \alpha_4 = 0$ and $\alpha_4 = 0$, that is, $\alpha_4 = \alpha_3 = \alpha_2 = \alpha_1 = 0$.

How to Decide Linear Independence of Column Vectors?

Suppose $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are column vectors each of length m. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be the matrix having $\mathbf{c}_1, \ldots, \mathbf{c}_n$ as its n columns. Then for $x_1, \ldots, x_n \in \mathbb{R}$,

$$x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{A}\mathbf{x}.$$

Hence the subset $S := \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ of $\mathbb{R}^{m \times 1}$ is linearly independent if and only if the linear system $\mathbf{A}\mathbf{x} = \mathbf{0}$ has only the zero solution. This is the case if and only if in any REF $\mathbf{A}'$ of $\mathbf{A}$, there are n nonzero rows, as we have seen in Lecture 3.

Hence if $\mathbf{A} \in \mathbb{R}^{m \times n}$ is transformed to a REF $\mathbf{A}'$, and $\mathbf{A}'$ has r nonzero rows, then the columns of $\mathbf{A}$ form a linearly independent subset of $\mathbb{R}^{m \times 1}$ if $r = n$, and they form a linearly dependent subset $\mathbb{R}^{m \times 1}$ if $r < n$.

Row Rank of a Matrix

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. The **row rank** of $\mathbf{A}$ is the maximum number of linearly independent row vectors of $\mathbf{A}$. Thus the row rank of $\mathbf{A}$ is equal to r if and only if there is a linearly independent set of r rows of $\mathbf{A}$ and any set of $r + 1$ rows of $\mathbf{A}$ is linearly dependent.

Let r be the row rank of $\mathbf{A}$. Then $r = 0$ if and only if $\mathbf{A} = \mathbf{O}$. Since the total number of rows of $\mathbf{A}$ is m, we see that $r \leq m$. Also, since the row vectors of $\mathbf{A}$ form a subset of $\mathbb{R}^{1 \times n}$, no more than n of them can be linearly independent. Thus $r \leq n$.

Let $\mathbf{a}_1, \ldots, \mathbf{a}_m$ be any m vectors in $\mathbb{R}^{1 \times n}$. Clearly, they are linearly independent if and only if the matrix $\mathbf{A}$ formed with these vectors as row vectors has row rank m, and they are linearly dependent if the row rank of $\mathbf{A}$ is less than m.

Examples

(i) Let $\mathbf{A} := \begin{bmatrix} 3 & 0 & 2 & 2 \\ -3 & 0 & 12 & 27 \\ -21 & 21 & 0 & 15 \end{bmatrix}$.

The row vectors of $\mathbf{A}$ are $\mathbf{a}_1 := \begin{bmatrix} 3 & 0 & 2 & 2 \end{bmatrix}$, $\mathbf{a}_2 := \begin{bmatrix} -3 & 0 & 12 & 27 \end{bmatrix}$ and $\mathbf{a}_3 := \begin{bmatrix} -21 & 21 & 0 & 15 \end{bmatrix}$.

Let $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ be such that $\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \alpha_3\mathbf{a}_3 = \mathbf{0}$.

This means

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} \begin{bmatrix} 3 & 0 & 2 & 2 \\ -3 & 0 & 12 & 27 \\ -21 & 21 & 0 & 15 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix},$$

that is, $3\alpha_1 - 3\alpha_2 - 21\alpha_3 = 0$, $21\alpha_3 = 0, 2\alpha_1 + 12\alpha_2 = 0$ and $2\alpha_1 + 27\alpha_2 + 15\alpha_3 = 0$. Clearly, $\alpha_3 = 0$, and the two equations $3\alpha_1 - 3\alpha_2 = 0, 2\alpha_1 + 12\alpha_2 = 0$ show that $\alpha_1 = \alpha_2 = 0$ as well. Thus the 3 rows of $\mathbf{A}$ are linearly independent. Hence the row rank of $\mathbf{A}$ is 3.

(ii)

Let $\mathbf{A} := \begin{bmatrix} 3 & 0 & 2 & 2 \\ -3 & 21 & 12 & 27 \\ -21 & 21 & 0 & 15 \end{bmatrix}$.

The row vectors of $\mathbf{A}$ are $\mathbf{a}_1 := \begin{bmatrix} 3 & 0 & 2 & 2 \end{bmatrix}$, $\mathbf{a}_2 := \begin{bmatrix} -3 & 21 & 12 & 27 \end{bmatrix}$ and $\mathbf{a}_3 := \begin{bmatrix} -21 & 21 & 0 & 15 \end{bmatrix}$. Observe that $6\mathbf{a}_1 - \mathbf{a}_2 + \mathbf{a}_3 = \mathbf{0}$. Hence the three row vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are not linearly independent. But the set $\{\mathbf{a}_1, \mathbf{a}_2\}$ is linearly independent since $\mathbf{a}_1 \neq \mathbf{0}, \mathbf{a}_2 \neq \mathbf{0}$, and they are not scalar multiples of each other. (The same holds for the sets $\{\mathbf{a}_2, \mathbf{a}_3\}$ and $\{\mathbf{a}_3, \mathbf{a}_1\}$.) Hence the row rank of $\mathbf{A}$ is 2.

We used the relation $6\mathbf{a}_1 - \mathbf{a}_2 + \mathbf{a}_3 = \mathbf{0}$ above to determine the row rank of $\mathbf{A}$. It is difficult to think of such a relation out of nowhere. We shall develop a systematic approach to find the row rank of a matrix.

First we prove the following preliminary results.

Proposition

If a matrix $\mathbf{A}$ is transformed to a matrix $\mathbf{A}'$ by elementary row operations, then the row ranks of $\mathbf{A}$ and $\mathbf{A}'$ are equal, that is, EROs do not alter the row rank of a matrix.

Proof. ERO of type I: $R_i \longleftrightarrow R_j$ with $i \neq j$: $\mathbf{A}$ and $\mathbf{A}'$ have the same set of row vectors. So there is nothing to prove.

ERO of type II: $R_i + \alpha R_j$ with $i \neq j$: Suppose the set $\{\mathbf{a}_1, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_m\}$ of all row vectors of $\mathbf{A}$ contains a linearly independent subset $S := \{\mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_s}\}$ having s elements. We claim that the set $\{\mathbf{a}_1, \ldots, \mathbf{a}_i + \alpha\, \mathbf{a}_j, \ldots, \mathbf{a}_j, \ldots \mathbf{a}_m\}$ of all row vectors of $\mathbf{A}'$ also contains a linearly independent set S' containing s elements. If $\mathbf{a}_i \notin S$, then we let $S' := S$. Next, suppose $\mathbf{a}_i \in S$. Then we may replace $\mathbf{a}_i$ suitably either by $\mathbf{a}_i + \alpha \mathbf{a}_j$ or by $\mathbf{a}_j$ in the set S to form a linearly independent set S'. The last statement follows by considering the cases $\mathbf{a}_i + \alpha \mathbf{a}_j = \mathbf{0}$, $\mathbf{a}_j = \mathbf{0}$, and by observing that if $\mathbf{a}_i + \alpha\, \mathbf{a}_j$ as well as $\mathbf{a}_j$ were linear combinations of vectors in $S \setminus \{\mathbf{a}_i\}$, then so would be $\mathbf{a}_i = (\mathbf{a}_i + \alpha\, \mathbf{a}_j) - \alpha\, \mathbf{a}_j$, and this would contradict the linear independence of S. Note that the converse claim also holds.

ERO of type III: $\alpha\, R_j$ with $\alpha \neq 0$: $\{\mathbf{a}_j, \mathbf{a}_{j_1}, \ldots \mathbf{a}_{j_s}\}$ is linearly independent $\iff$ $\{\alpha\, \mathbf{a}_j, \mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_s}\}$ is linearly independent.

Thus the maximum number of linearly independent rows of $\mathbf{A}$ is the same as the maximum number of linearly independent rows of $\mathbf{A}'$, that is, the row ranks of $\mathbf{A}$ and $\mathbf{A}'$ are equal. $\square$

Proposition

Let a matrix $\mathbf{A}'$ be in REF. Then the nonzero rows of $\mathbf{A}'$ are linearly independent, and so the row rank of $\mathbf{A}'$ is equal to the number of nonzero rows of $\mathbf{A}'$.

Proof. Let the number of nonzero rows of $\mathbf{A}'$ be r. Let the pivots $p_1, \ldots, p_r$ in these rows be in columns $k_1, \ldots, k_r$, where $1 \leq k_1 < \cdots < k_r \leq n$. Suppose $\alpha_1 \mathbf{a}_1 + \cdots + \alpha_r \mathbf{a}_r = \mathbf{0}$, where $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$.

Assume for a moment that not all $\alpha_1, \ldots, \alpha_r$ are zero, and let α_j be the first nonzero number among them. Now the k_jth component of $\alpha_1 \mathbf{a}_1 + \cdots + \alpha_r \mathbf{a}_r = \alpha_j \mathbf{a}_j + \cdots + \alpha_r \mathbf{a}_r$ is equal to $\alpha_j p_j$ since all entries in the k_jth column below the pivot p_j are equal to 0. Hence $\alpha_j p_j = 0$. But $p_j \neq 0$, and so $\alpha_j = 0$, contrary to our assumption. Thus $\alpha_1 = \cdots = \alpha_r = 0$. This shows that the first r rows of $\mathbf{A}'$ are linearly independent.

Also, since the last $m - r$ rows of $\mathbf{A}'$ are zero rows, any $r + 1$ row vectors of $\mathbf{A}'$ will contain the vector $\mathbf{0}$, and so they will not be linearly independent. Hence the row rank of $\mathbf{A}'$ is r. $\square$

We have now obtained an important result which tells us how to find the row rank of a matrix.

Proposition

The row rank of a matrix is equal to the number of nonzero rows in any row echelon form of the matrix.

Proof. Let $\mathbf{A}$ be a $m \times n$ matrix. By using EROs of type I and type II, we transform $\mathbf{A}$ to a row echelon form $\mathbf{A}'$. Then the row rank of $\mathbf{A}$ is equal to the row rank of $\mathbf{A}'$, and it is equal to the number of nonzero rows of $\mathbf{A}'$. $\square$

The above proposition implies that if $\mathbf{A}'$ and $\mathbf{A}''$ are two row echelon forms of a matrix $\mathbf{A}$, then they have the same number of nonzero rows; this number is equal to the row rank of $\mathbf{A}$.

Since each nonzero row of a matrix in a REF has exactly one pivot, we see that the row rank of a matrix is equal to the number of pivots in any row echelon form of the matrix.