MA 110
Linear Algebra and Differential Equations
Lecture 03

Prof. Sudhir R. Ghorpade
Department of Mathematics
IIT Bombay
http://www.math.iitb.ac.in/∼srg/

Spring 2025

Recall: Let **A** be an $m \times n$ matrix. We say that **A** is in a row echelon form (REF) if

(i) The nonzero rows of **A** precede the zero rows of **A**.

(ii) If **A** has r nonzero rows, where $r \in \mathbb{N}$, and the pivot in row 1 appears in the column k_1, the pivot in row 2 appears in the column k_2, and so on the pivot in row r appears in the column k_r, then $k_1 < k_2 < \cdots < k_r$.

For example, the matrix $\begin{bmatrix} \boxed{5} & 6 & 7 & 8 \\ 0 & \boxed{1} & 4 & 2 \\ 0 & 0 & 0 & \boxed{3} \end{bmatrix}$ is in a REF.

Note: In French, 'échelon' means 'level'.

Summary of results discussed in Lecture 2:

Suppose an $m \times n$ matrix $\mathbf{A}$ is in a REF, and let r be the number of nonzero rows of $\mathbf{A}$. If $\mathbf{b} \in \mathbb{R}^{m \times 1}$, then the linear system $\mathbf{Ax} = \mathbf{b}$ has

(i) no solution if one of $b_{r+1}, \ldots, b_m$ is nonzero.

(ii) a unique solution if $b_{r+1} = \cdots = b_m = 0$ and $r = n$.

(iii) infinitely many solutions if $b_{r+1} = \cdots = b_m = 0$ and $r < n$. (In this case, there are $n - r$ free variables which give $n - r$ degrees of freedom.)

Considering the case $\mathbf{b} = \mathbf{0} \in \mathbb{R}^{m \times 1}$ and recalling that $r \leq m$, we obtain the following important results.

Proposition

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be in REF with r nonzero rows. Then the linear system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution if and only if $r = n$. In particular, if $m < n$, then $\mathbf{Ax} = \mathbf{0}$ has a nonzero solution.

Gauss Elimination Method (GEM)

We have seen how to solve the linear system $\mathbf{Ax} = \mathbf{b}$ when the matrix $\mathbf{A}$ is in a row echelon form (REF).

We now explain the **Gauss Elimination Method** (GEM) by which we can transform any $\mathbf{A} \in \mathbb{R}^{m \times n}$ to a REF.
This involves the following two **elementary row operations** (EROs):

Type I: Interchange of two rows

Type II: Addition of a scalar multiple of a row to another row

We shall later consider the following elementary row operation:

Type III: Multiplication of a row by a nonzero scalar

First we remark that if the augmented matrix $[\mathbf{A}|\mathbf{b}]$ is transformed to a matrix $[\mathbf{A}'|\mathbf{b}']$ by any of the EROs, then $\mathbf{A}\mathbf{x} = \mathbf{b} \iff \mathbf{A}'\mathbf{x} = \mathbf{b}'$ for $\mathbf{x} \in \mathbb{R}^{n \times 1}$, that is, the linear systems $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ have the same solutions.

This follows by noting that an interchange of two equations does not change the solutions, neither does an addition of an equation to another, nor does a multiplication of an equation by a nonzero scalar, since these operations can be undone by similar operations, namely, interchange of the equations in the reverse order, subtraction of an equation from another, and division of an equation by a nonzero scalar.

Consequently, we are justified in performing EROs on the augmented matrix $[\mathbf{A}|\mathbf{b}]$ in order to obtain all solutions of the given linear system.

Transformation to REF

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$, that is, let $\mathbf{A}$ be an $m \times n$ matrix with entries in $\mathbb{R}$. If $\mathbf{A} = \mathbf{O}$, the zero matrix, then it is already in REF.

Suppose $\mathbf{A} \neq \mathbf{O}$.

(i) Let column k_1 be the first nonzero column of $\mathbf{A}$, and let some nonzero entry p_1 in this column occur in the jth row of $\mathbf{A}$. Interchange row j and row 1. Then $\mathbf{A}$ is transformed to

$$\mathbf{A}' := \begin{bmatrix} 0 & \cdots & 0 & \boxed{p_1} & * & \cdots & * \\ 0 & \cdots & 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & * & * & \cdots & * \end{bmatrix},$$

where $*$ denotes a real number. Note: p_1 becomes the chosen pivot in row 1. (This choice may not be unique.)

(ii) Since $p_1 \neq 0$, add suitable scalar multiples of row 1 of $\mathbf{A}'$ to rows 2 to m of $\mathbf{A}'$, so that all entries in column k_1 below the pivot p_1 are equal to 0. Then $\mathbf{A}'$ is transformed to

$$\mathbf{A}'' := \begin{bmatrix} 0 & \cdots & 0 & \boxed{p_1} & * & \cdots & * \\ 0 & \cdots & 0 & \boxed{0} & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \boxed{0} & * & \cdots & * \end{bmatrix}.$$

(iii) Keep row 1 of $\mathbf{A}''$ intact, and repeat the above process for the remaining $(m-1) \times n$ submatrix of $\mathbf{A}''$ to obtain

$$\mathbf{A}''' := \begin{bmatrix} 0 & \cdots & 0 & \boxed{p_1} & * & * & * & * & \cdots & * \\ 0 & \cdots & 0 & \boxed{0} & \cdots & 0 & \boxed{p_2} & * & \cdots & * \\ 0 & \cdots & 0 & \boxed{0} & \cdots & 0 & \boxed{0} & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \boxed{0} & \cdots & 0 & \boxed{0} & * & \cdots & * \end{bmatrix},$$

where $p_2 \neq 0$ and occurs in column k_2 of $\mathbf{A}'''$, where $k_1 < k_2$.

Note: p_2 becomes the chosen pivot in row 2. (Again, this choice may not be unique.)

(iv) Keep rows 1 and 2 of $\mathbf{A}'''$ intact, and repeat the above process till the remaining submatrix has no nonzero row. The resulting $m \times n$ matrix is in REF with pivots $p_1, \ldots, p_r$ in columns $k_1, \ldots, k_r$, and the last $m - r$ rows are zero rows, where $1 \leq r \leq m$.

Notation

$R_i \longleftrightarrow R_j$ will denote the interchange of the ith row R_i and the jth row R_j for $1 \leq i, j \leq m$ with $i \neq j$.

$R_i + \alpha R_j$ will denote the addition of α times the jth row R_j to the ith row R_i for $1 \leq i, j \leq m$ with $i \neq j$.

αR_j will denote the multiplication of the jth row R_j by the nonzero scalar α for $1 \leq j \leq m$.

Remark

A matrix $\mathbf{A}$ may be transformed to different REFs by EROs.

For example, we can transform $\mathbf{A} := \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$ by EROs to $\begin{bmatrix} 1 & 3 \\ 0 & -6 \end{bmatrix}$ as well as to $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, both of which are REFs.

Recommendation

While interchanging rows, choose as large a pivot as possible. (In the above example, the pivot 2 is larger than the pivot 1.) This strategy is known as **partial pivoting**.

Example Let $\mathbf{A} := \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} := \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Then $R_1 \longleftrightarrow R_2$ gives $\mathbf{x} := \begin{bmatrix} 1 & 1 \end{bmatrix}^{\mathsf{T}}$ as the solution of the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Let $\epsilon > 0$ with $\epsilon \neq 1$ and $\mathbf{A}_\epsilon := \begin{bmatrix} \epsilon & 1 \\ 1 & 1 \end{bmatrix}$. Try solving the linear system $\mathbf{A}_\epsilon \mathbf{x} = \mathbf{b}$ on a computer with $\epsilon = 10^{-4}, 10^{-5}, 10^{-6}$.

Examples

(i) Consider the linear system

$$\begin{aligned} 3x_1 + 2x_2 + \ x_3 &= 3 \\ 2x_1 + \ x_2 + \ x_3 &= 0 \\ 6x_1 + 2x_2 + 4x_3 &= 6. \end{aligned}$$

We can check that

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 2 & 1 & 1 & 0 \\ 6 & 2 & 4 & 6 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} \boxed{3} & 2 & 1 & 3 \\ 0 & \boxed{-1/3} & 1/3 & -2 \\ 0 & -2 & 2 & 0 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|c} \boxed{3} & 2 & 1 & 3 \\ 0 & \boxed{-1/3} & 1/3 & -2 \\ 0 & 0 & 0 & 12 \end{array}\right].$$

Here $m = 3 = n$, $r = 2$ and $b'_{r+1} = b'_3 = 12 \neq 0$. Hence the given linear system has no solution.

(ii) Consider the linear system

$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ -x_1 + x_2 - x_3 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ 20x_1 + 10x_2 &= 80. \end{aligned}$$

As we have already seen in Lecture 2,

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 10 & 25 & 90 \\ 20 & 10 & 0 & 80 \end{array}\right] \xrightarrow{\text{EROs}} \left[\begin{array}{ccc|c} \boxed{1} & -1 & 1 & 0 \\ 0 & \boxed{10} & 25 & 90 \\ 0 & 0 & \boxed{-95} & -190 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Here $m = 4$, $n = 3$, $r = 3$, pivotal columns: $\boxed{1}$, $\boxed{2}$, $\boxed{3}$.

Since $b'_{r+1} = b'_4 = 0$ and $r = n$, the linear system has a unique solution, namely $\mathbf{x}_0 := \begin{bmatrix} 2 & 4 & 2 \end{bmatrix}^\mathsf{T}$, which we had obtained by back substitution in Lecture 2.

(iii) Consider the linear system

$$\begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ 0.6x_1 + 1.5x_2 + 1.5x_3 - 5.4x_4 &= 2.7 \\ 1.2x_1 - 0.3x_2 - 0.3x_3 + 2.4x_4 &= 2.1. \end{aligned}$$

We can check that

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{cccc|c} 3 & 2 & 2 & -5 & 8 \\ 0.6 & 1.5 & 1.5 & -5.4 & 2.7 \\ 1.2 & -0.3 & -0.3 & 2.4 & 2.1 \end{array}\right]$$

$$\xrightarrow{\text{EROs}} \left[\begin{array}{cccc|c} \boxed{3} & 2 & 2 & -5 & 8 \\ 0 & \boxed{1.1} & 1.1 & -4.4 & 1.1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] =: [\mathbf{A}'|\mathbf{b}'].$$

Here $m = 3, n = 4, r = 2$, pivotal columns: $\boxed{1}$, $\boxed{2}$,
nonpivotal columns: $\boxed{3}$, $\boxed{4}$.

Since $b'_{r+1} = b'_3 = 0$, the linear system $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ has a solution.

For a particular solution of $\mathbf{A}'\mathbf{x} = \mathbf{b}'$, let $x_3 = x_4 = 0$. Then
$1.1\,x_2 = 1.1 \implies x_2 = 1$,
$3x_1 + 2x_2 = 8 \implies x_1 = 2$,
Thus $\mathbf{x}_0 := \begin{bmatrix} 2 & 1 & 0 & 0 \end{bmatrix}^{\mathsf{T}}$ is a particular solution.

Since $r = 2 < 4 = n$, the linear system has many solutions.
For basic solutions of $\mathbf{A}'\mathbf{x} = \mathbf{0}'$, where $\mathbf{0}' = \mathbf{0}$,
let $x_3 = 1, x_4 = 0$, so that $\mathbf{s}_3 := \begin{bmatrix} 0 & -1 & 1 & 0 \end{bmatrix}^{\mathsf{T}}$,
and $x_4 = 1, x_3 = 0$, so that $\mathbf{s}_4 := \begin{bmatrix} -1 & 4 & 0 & 1 \end{bmatrix}^{\mathsf{T}}$,

The general solution of $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ is given by
$\mathbf{x} = \mathbf{x}_0 + \alpha_3\mathbf{s}_3 + \alpha_4\mathbf{s}_4$, that is,
$x_1 = 2 - \alpha_4,\; x_2 = 1 - \alpha_3 + 4\alpha_4,\; x_3 = \alpha_3,\; x_4 = \alpha_4$, where
α_3, α_4 are arbitrary real numbers. These are precisely the solutions of the given linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Proposition

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then the linear system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution if and only if any REF of $\mathbf{A}$ has n nonzero rows. In particular, if $m < n$, then $\mathbf{Ax} = \mathbf{0}$ has a nonzero solution.

Proof. We saw that these results hold if $\mathbf{A}$ itself is in REF. Since every $\mathbf{A} \in \mathbb{R}^{m \times n}$ can be transformed to a REF $\mathbf{A}'$ by EROs, and since the solutions of the linear system $\mathbf{Ax} = \mathbf{0}$ and the transformed system $\mathbf{A}'\mathbf{x} = \mathbf{0}'$, where $\mathbf{0}' = \mathbf{0}$, are the same, the desired results hold. □

Note: Suppose an $m \times n$ matrix $\mathbf{A}$ is transformed by EROs to different REFs $\mathbf{A}'$ and $\mathbf{A}''$. Suppose $\mathbf{A}'$ has r' nonzero rows and $\mathbf{A}''$ has r'' nonzero rows. Then $0 \leq r', r'' \leq \min\{m, n\}$. The above result implies that $r' = n \iff r'' = n$. We shall later see that $r' = r''$ always.

Inverse of a Square Matrix

We now introduce a special kind of square matrices.

Let $\mathbf{A}$ be a square matrix of size $n \in \mathbb{N}$, that is, let $\mathbf{A} \in \mathbb{R}^{n \times n}$. We say that $\mathbf{A}$ is **invertible** if there is $\mathbf{B} \in \mathbb{R}^{n \times n}$ such that $\mathbf{AB} = \mathbf{I} = \mathbf{BA}$, and in this case, $\mathbf{B}$ is called an **inverse** of $\mathbf{A}$.

Examples

The matrix $\mathbf{A} := \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ is invertible. To see this, let $\mathbf{B} := \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}$, and check $\mathbf{AB} = \mathbf{I} = \mathbf{BA}$. On the other hand, the nonzero matrix $\mathbf{A} := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is not invertible. To see this, let $\mathbf{B} := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ and note that $\mathbf{AB} = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \neq \mathbf{I}$.

If $\mathbf{A}$ is invertible, then it has a unique inverse. In fact, if $\mathbf{AC} = \mathbf{I} = \mathbf{BA}$, then $\mathbf{C} = \mathbf{IC} = (\mathbf{BA})\mathbf{C} = \mathbf{B}(\mathbf{AC}) = \mathbf{BI} = \mathbf{B}$ by the associativity of the matrix multiplication.

If $\mathbf{A}$ is invertible, its inverse will be denoted by $\mathbf{A}^{-1}$, and so $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$. Clearly, $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$. If $\mathbf{A}$ is invertible and if one can guess its inverse, then it is easy to verify that it is in fact the inverse of $\mathbf{A}$. Here is a case in point.

Proposition

Let $\mathbf{A}$ be a square matrix. Then $\mathbf{A}$ is invertible if and only if $\mathbf{A}^\mathsf{T}$ is invertible. In this case, $(\mathbf{A}^\mathsf{T})^{-\mathbf{1}} = (\mathbf{A}^{-\mathbf{1}})^\mathsf{T}$.

Proof. Suppose $\mathbf{A}$ is invertible and $\mathbf{B}$ is its inverse. Then $\mathbf{AB} = \mathbf{I} = \mathbf{BA}$, and so $\mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T} = \mathbf{I}^\mathsf{T} = \mathbf{A}^\mathsf{T}\mathbf{B}^\mathsf{T}$. Since $\mathbf{I}^\mathsf{T} = \mathbf{I}$, we see that $\mathbf{A}^\mathsf{T}$ is invertible and $(\mathbf{A}^\mathsf{T})^{-\mathbf{1}} = \mathbf{B}^\mathsf{T} = (\mathbf{A}^{-\mathbf{1}})^\mathsf{T}$.

Next, if $\mathbf{A}^\mathsf{T}$ is invertible, then $\mathbf{A} = (\mathbf{A}^\mathsf{T})^\mathsf{T}$ is invertible. □

We now relate the invertibility of a square matrix $\mathbf{A}$ to the solutions of the homogeneous system $\mathbf{Ax} = \mathbf{0}$.

Proposition

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$. Then $\mathbf{A}$ is invertible if and only if the linear system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution.

Proof. Suppose $\mathbf{A}$ is invertible. Then by definition, there is $\mathbf{B} \in \mathbb{R}^{n \times n}$ such that $\mathbf{BA} = \mathbf{I}$. If $\mathbf{x} \in \mathbb{R}^{n \times 1}$ satisfies $\mathbf{Ax} = \mathbf{0}$, then $\mathbf{x} = \mathbf{BAx} = \mathbf{B}(\mathbf{Ax}) = \mathbf{B0} = \mathbf{0}$. Thus the linear system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution.

Conversely, suppose the linear system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution. Let $\mathbf{y} = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{n \times 1}$. We transform the augmented matrix $[\mathbf{A}|\mathbf{y}]$ to a matrix $[\mathbf{A}'|\mathbf{y}']$, where $\mathbf{A}'$ is in REF. By our previous result, $\mathbf{A}'$ has n nonzero rows, and so back substitution gives a unique $\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{n \times 1}$ such that $\mathbf{A}'\mathbf{x} = \mathbf{y}'$. Hence $\mathbf{Ax} = \mathbf{y}$.

Further, the process of the back substitution shows that the entries $x_1, \ldots, x_n$ of $\mathbf{x}$ are given as follows:

$$\begin{array}{rcl} x_n & = & c'_{nn} y'_n \\ x_{n-1} & = & c'_{(n-1)(n-1)} y'_{n-1} + c'_{(n-1)n} y'_n \\ \vdots & \vdots & \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ x_2 & = & c'_{22} y'_2 + \cdots + \cdots + \cdots + c'_{2n} y'_n \\ x_1 & = & c'_{11} y'_1 + c'_{12} y'_2 + \cdots + \cdots + \cdots + c'_{1n} y'_n, \end{array}$$

where $\mathbf{y}' = \begin{bmatrix} y'_1 & \cdots & y'_n \end{bmatrix}^{\mathsf{T}}$ and $c'_{jk} \in \mathbb{R}$ for $j, k = 1, \ldots, n$.

Also, since $\mathbf{y}'$ is obtained from $\mathbf{y}$ by performing EROs (which are of the type $R_i \longleftrightarrow R_j$, $R_i + \alpha R_j$ and αR_j) on $[\mathbf{A}|\mathbf{y}]$, we see that each $y'_1, \ldots, y'_n$ is a linear combination of the entries $y_1, \ldots, y_n$ of $\mathbf{y}$. As a result, each $x_1, \ldots, x_n$ is a linear combination of $y_1, \ldots, y_n$.

Thus there is $c_{jk} \in \mathbb{R}$ for $j, k = 1, \ldots, n$ (not depending on $y_1, \ldots, y_n$) such that

$$\begin{aligned} x_1 &= c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n \\ x_2 &= c_{21}y_1 + c_{22}y_2 + \cdots + c_{2n}y_n \\ \vdots \; &\vdots \; \vdots \qquad \vdots \qquad\qquad \vdots \\ x_n &= c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n. \end{aligned}$$

Define $\mathbf{C} := \begin{bmatrix} c_{jk} \end{bmatrix} \in \mathbb{R}^{n\times n}$. Then $\mathbf{x} = \mathbf{Cy}$, and so $\mathbf{ACy} = \mathbf{A}(\mathbf{Cy}) = \mathbf{Ax} = \mathbf{y}$. Letting $\mathbf{y} := \mathbf{e}_k \in \mathbb{R}^{n\times 1}$, we see that $(\mathbf{AC})\mathbf{e_k} = \mathbf{e_k}$ for $k = 1, \ldots, n$. Hence $\mathbf{AC} = \mathbf{I}$. If $\mathbf{y} \in \mathbb{R}^{n\times 1}$ satisfies $\mathbf{Cy} = \mathbf{0}$, then $\mathbf{y} = \mathbf{ACy} = \mathbf{A}(\mathbf{Cy}) = \mathbf{A0} = \mathbf{0}$. Thus the linear system $\mathbf{Cy} = \mathbf{0}$ has only the zero solution.

By what we have proved above, there is $\mathbf{D} \in \mathbb{R}^{n\times n}$ such that $\mathbf{CD} = \mathbf{I}$. In fact, $\mathbf{D} = \mathbf{ID} = (\mathbf{AC})\mathbf{D} = \mathbf{A}(\mathbf{CD}) = \mathbf{AI} = \mathbf{A}$ by the associativity of the matrix multiplication.

Thus $\mathbf{AC} = \mathbf{I} = \mathbf{CA}$, and so $\mathbf{A}$ is invertible. □

Corollary

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$. If there is $\mathbf{B} \in \mathbb{R}^{n\times n}$ such that either $\mathbf{BA} = \mathbf{I}$ or $\mathbf{AB} = \mathbf{I}$, then $\mathbf{A}$ is invertible, and $\mathbf{A}^{-1} = \mathbf{B}$.

Proof. Let $\mathbf{B} \in \mathbb{R}^{n\times n}$ be such that $\mathbf{BA} = \mathbf{I}$. If $\mathbf{x} \in \mathbb{R}^{n\times 1}$ satisfies $\mathbf{Ax} = \mathbf{0}$, then $\mathbf{x} = \mathbf{BAx} = \mathbf{B}(\mathbf{Ax}) = \mathbf{B0} = \mathbf{0}$. Thus the linear system $\mathbf{Ax} = \mathbf{0}$ has only the zero solution. By the previous proposition, $\mathbf{A}$ is invertible. Then there is $\mathbf{C} \in \mathbb{R}^{\mathbf{n}\times\mathbf{n}}$ such that $\mathbf{AC} = \mathbf{I}$, and $\mathbf{B} = \mathbf{C}$. Hence $\mathbf{A}^{-1} = \mathbf{B}$.

Next, let $\mathbf{B} \in \mathbb{R}^{n\times n}$ be such that $\mathbf{AB} = \mathbf{I}$. Then $\mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}} = \mathbf{I}^{\mathsf{T}} = \mathbf{I}$. By what we have just proved, $\mathbf{A}^{\mathsf{T}}$ is invertible, and $(\mathbf{A}^{\mathsf{T}})^{-1} = \mathbf{B}^{\mathsf{T}}$. Hence $\mathbf{A} = (\mathbf{A}^{\mathsf{T}})^{\mathsf{T}}$ is invertible, and $\mathbf{A}^{-1} = (\mathbf{B}^{\mathsf{T}})^{\mathsf{T}} = \mathbf{B}$. □

Note: The above result is a definite improvement over requiring the existence of a matrix $\mathbf{B}$ satisfying both $\mathbf{BA} = \mathbf{I}$ and $\mathbf{AB} = \mathbf{I}$ for the invertibility of a square matrix $\mathbf{A}$.

Proposition

Let $\mathbf{A}$ and $\mathbf{B}$ be square matrices. Then $\mathbf{AB}$ is invertible if and only if $\mathbf{A}$ and $\mathbf{B}$ are invertible, and then $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Proof. Let $\mathbf{A}$ and $\mathbf{B}$ be invertible. Using the associativity of matrix multiplication, we easily see that

$$(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) = \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}.$$

Hence $\mathbf{AB}$ is invertible and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ by the previous corollary.

Conversely, let $\mathbf{AB}$ be invertible. Then there is $\mathbf{C}$ such that $(\mathbf{AB})\mathbf{C} = \mathbf{I} = \mathbf{C}(\mathbf{AB})$. Since $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} = \mathbf{I}$, we see that $\mathbf{A}$ is invertible, and $\mathbf{A}^{-1} = \mathbf{BC}$. Also, since $(\mathbf{CA})\mathbf{B} = \mathbf{C}(\mathbf{AB}) = \mathbf{I}$, we see that $\mathbf{B}$ is invertible and $\mathbf{B}^{-1} = \mathbf{CA}$, again by the previous corollary. □