MA 110
Linear Algebra and Differential Equations
Lecture 04

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

Spring 2025

Recall: A square matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{n}\times\mathbf{n}}$ is said to be 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 $\mathbf{A}$.

We have seen examples of square matrices that are invertible and also those that are not invertible. Further we noted that:

- If a square matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{n}\times\mathbf{n}}$ is invertible, then it has a unique inverse, and it is denoted by $\mathbf{A}^{-1}$
- If a square matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{n}\times\mathbf{n}}$ is invertible, then so is its transpose $\mathbf{A}^{\mathsf{T}}$ and in this case,

$$(\mathbf{A}^{\mathsf{T}})^{-\mathbf{1}} = (\mathbf{A}^{-\mathbf{1}})^{\mathsf{T}}.$$

We have also related the invertibility of a square matrix $\mathbf{A}$ to the solutions of the homogeneous system $\mathbf{Ax} = \mathbf{0}$ by proving:

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.

This gave a 50% reduction in the condition for invertibility:

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 $\mathbf{A}^{-1} = \mathbf{B}$.

And we also established the following useful property:

Proposition

Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$. Then $\mathbf{AB}$ is invertible if and only if both $\mathbf{A}$ and $\mathbf{B}$ are invertible,. In this case, $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Row Canonical Form (RCF)

As we have seen, a matrix **A** may not have a unique REF. However, a special REF of **A** turns out to be unique.

An $m \times n$ matrix **A** is said to be in a **row canonical form** (RCF) or a **reduced row echelon form** (RREF) if
(i) it is in a row echelon form (REF),
(ii) all pivots are equal to 1 and
(iii) in each pivotal column, all entries above the pivot are (also) equal to 0.

For example, the matrix

$$\mathbf{A} := \begin{bmatrix} 0 & \boxed{1} & * & 0 & 0 & * \\ 0 & 0 & 0 & \boxed{1} & 0 & * \\ 0 & 0 & 0 & 0 & \boxed{1} & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

is in a RCF, where $*$ denotes any real number.

Note: If $\mathbf{A}$ is in REF, then in each pivotal column, all entries below the pivot are 0. If $\mathbf{A}$ is in fact in RCF and has r nonzero rows, then the $r \times r$ submatrix formed by the first r rows and the r pivotal columns is the $r \times r$ identity matrix $\mathbf{I}$.

Suppose an $m \times n$ matrix $\mathbf{A}$ is in RCF and has r nonzero rows. If $r = n$, then it has n pivotal columns, that is, all its columns are pivotal, and so $\mathbf{A} = \mathbf{I}$ if $m = n$, and $\mathbf{A} = \begin{bmatrix} \mathbf{I} \\ \\ \mathbf{O} \end{bmatrix}$ if $m > n$, where $\mathbf{I}$ is the $n \times n$ identity matrix and $\mathbf{O}$ is the $(m - n) \times n$ zero matrix.

To transform an $m \times n$ matrix to a RCF, we first transform it to a REF by elementary row operations of type I and II. Then we multiply a row containing a pivot p by $1/p$ (which is an elementary row operation of type III), and then we add a suitable nonzero multiple of this row to each preceding row.

Every matrix has a unique RCF. (Proof by induction on n)

Example

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 16 \end{bmatrix} \xrightarrow{\text{EROs}} \begin{bmatrix} \boxed{1} & 3 & -2 & 0 & 2 & 0 \\ 0 & 0 & \boxed{-1} & -2 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & \boxed{6} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

which is in REF,

$$\xrightarrow{\text{EROs}} \begin{bmatrix} \boxed{1} & 3 & 0 & 4 & 2 & 6 \\ 0 & 0 & \boxed{1} & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & \boxed{6} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{EROs}} \begin{bmatrix} \boxed{1} & 3 & 0 & 4 & 2 & 0 \\ 0 & 0 & \boxed{1} & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

which is in RCF.

Recall: A square matrix $\mathbf{A}$ is invertible if and only if the homogeneous linear system $\mathbf{A}\mathbf{x} = \mathbf{0}$ has only the zero solution.

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 $\mathbf{A}\mathbf{x} = \mathbf{0}$ 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 $\mathbf{I}\mathbf{x} = \mathbf{0} \implies \mathbf{x} = \mathbf{0}$ for $\mathbf{x} \in \mathbb{R}^{n \times 1}$, we see that the linear system $\mathbf{A}\mathbf{x} = \mathbf{0}$ has only the zero solution. Hence $\mathbf{A}$ is invertible. □

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 $[\mathbf{A} \,|\, \mathbf{I}]$ to $[\mathbf{I} \,|\, \mathbf{X}]$, where $\mathbf{X} \in \mathbb{R}^{3\times 3}$.

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

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

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

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

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. □