Linear Algebra

Second Edition

Vivek Sahai • Vikas Bist

Alpha Science

Linear Algebra

Second Edition

Vivek Sahai
Vikas Bist

Alpha Science International Ltd.
Oxford, U.K.

Linear Algebra
Second Edition
264 pgs.

Vivek Sahai
Department of Mathematics and Astronomy
Lucknow University
Lucknow

Vikas Bist
Department of Mathematics
Punjab University
Chandigarh

First Reprint 2002
Second Reprint 2007
Third Reprint 2011
Second Edition 2013

ALPHA SCIENCE INTERNATIONAL LTD.
7200 The Quorum, Oxford Business Park North
Garsington Road, Oxford OX4 2JZ, U.K.

www.alphasci.com

Printed from the camera-ready copy provided by the Authors.

ISBN 978-1-84265-808-6

Printed in India

Preface

This book presents a concise and comprehensive introduction to the fundamentals of linear algebra. The subject is developed so that it is accessible to the students of varied backgrounds. A very basic algebra and elementary properties of matrices and determinants are prerequisites. However, almost all the material that is needed for the text is given in the first chapter with reasonable details to make the book as much self contained as possible. A fairly large number of examples are included at the end of each section for conceptual understanding. Examples are also used to illustrate the computational aspect and to explore further related results. Each section ends with a collection of exercises with different levels of difficulty that help the reader to check their understanding and to apply it to concrete problems.

Chapter 2 deals with the elementary concepts like vector space, subspace, basis and dimension. Although the emphasis here is on finite dimensional vector spaces, examples of infinite dimensional vector spaces are also included to highlight the differences between these two classes. Linear transformations, isomorphism theorems, the matrix of a linear transformation, linear forms and dual spaces are also discussed in this chapter.

Chapter 3 is analysis of a single linear operator on a finite dimensional vector space. Characterizations of diagonalizable and triangulable operators appear in this chapter. The concept of generalized eigenvectors is used to obtain an inductive procedure for constructing a Jordan basis for a triangulable linear operator. The last section is about the rational canonical form and again the approach is algorithmic.

Chapter 4 is about finite dimensional inner product spaces. This chapter treats with normal and self adjoint operators and the spectral decomposition. Also it deals with nonnegative operators and polar and singular value decomposition.

The last chapter is on bilinear forms and extends the results of inner product spaces to bilinear spaces. The main thrust of this chapter is to classify finite dimensional reflexive bilinear spaces.

VIVEK SAHAI
VIKAS BIST

Contents

CHAPTER 1

Preliminaries

The purpose of this chapter is to give a brief account of a number of useful concepts and facts which are required in the text. The reader may be familiar with the contents of this chapter. Nonetheless, this may serve as a short review, and an introduction to the basic notation.

1. Matrices

We denote by $K^{m\times n}$ the set of all matrices of size $m \times n$, m rows and n columns, whose entries are from the field K. If $A \in K^{m\times n}$, then we denote its (i,j)–th entry by $[A]_{ij}$. A matrix of size $n \times 1$ (respectively, $1 \times n$) is called a **column** (respectively, **row**) **vector** of length n. A square matrix is a matrix of size $n \times n$.

The sum of two matrices is allowed whenever both are of the same size; for $A, B \in K^{m\times n}$, $A + B \in K^{m\times n}$ and

$$[A+B]_{ij} = [A]_{ij} + [B]_{ij}, \quad i = 1,\ldots,m, \ j = 1,\ldots,n.$$

The product AB of matrices A and B is defined if the number of columns of A is equal to the number of rows of B; for $A \in K^{m\times q}$, $B \in K^{q\times n}$, $AB \in K^{m\times n}$ and

$$[AB]_{ij} = \sum_{k=1}^{q} [A]_{ik} [B]_{kj}, \quad i = 1,\ldots,m, \ j = 1,\ldots,n.$$

The matrix addition and the matrix multiplication are associative:

$$(A+B)+C = A+(B+C), \text{ for all } A, B, C \in K^{m\times n}$$

and

$$(AB)C = A(BC), \text{ for all } A \in K^{m\times p}, B \in K^{p\times q}, C \in K^{q\times n}.$$

But the matrix addition is commutative and the matrix multiplication is not. We also have distributive laws:

$$A(B+C) = AB + AC, \text{ for all } A \in K^{m\times q}, \ B, C \in K^{q\times n},$$

and

$$(A+B)C = AC + BC, \text{ for all } A, B \in K^{m\times q}, C \in K^{q\times n}.$$

The elements of the field K are called **scalars**. Multiplication of an $m \times n$ matrix A by a scalar λ is a matrix λA of size $m \times n$ such that

$$[\lambda A]_{ij} = \lambda [A]_{ij}, \; i = 1, \ldots, m, \; j = 1, \ldots, n.$$

The following are basic properties of scalar multiplication. The reader can easily verify them.

$$\begin{aligned} (\lambda + \mu) A &= \lambda A + \mu A, \\ \lambda (A + B) &= \lambda A + \lambda B, \\ (\lambda\mu) A &= \lambda (\mu A), \\ 1A &= A, \end{aligned}$$

for all $\lambda, \mu \in K$ and $A, B \in K^{m \times n}$.

A sum of the form $\sum_{i=1}^{r} \lambda_i A_i$, where $\lambda_1, \ldots, \lambda_r \in K$ and $A_1, \ldots, A_r \in K^{m \times n}$ is called a **linear combination** of matrices. If $A_1, \ldots, A_r$ are column (respectively, row) vectors, then we call it a linear combination of column (respectively, row) vectors.

Let A be an $m \times n$ matrix. A **submatrix** of A is a matrix obtained from A by deleting its some rows or some columns or both. A **partition** of A is its expression into smaller submatrices:

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rs} \end{bmatrix},$$

where each A_{ij} is an $m_i \times n_j$ submatrix of A, and $m_1 + \cdots + m_r = m$, $n_1 + \cdots + n_s = n$. A partitioned matrix is also called a **block matrix**.

For example, let

$$A = \begin{bmatrix} 1 & 3 & 2 & 9 & 2 & 5 & 3 \\ 3 & 9 & 7 & 8 & 3 & 6 & 7 \\ 2 & 0 & 0 & 1 & 3 & 2 & 7 \\ 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 7 & 8 & 7 & 8 & 9 & 3 & 1 \end{bmatrix}.$$

Then $\begin{bmatrix} 3 & 9 & 3 \\ 0 & 1 & 7 \\ 8 & 8 & 1 \end{bmatrix}$ is a submatrix of A obtained by removing rows 2 and 4, and columns 1, 3, 4 and 5.

$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$ is a partition of A, where

$$A_{11} = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 9 & 7 \\ 2 & 0 & 0 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 9 & 2 \\ 8 & 3 \\ 1 & 3 \end{bmatrix}, \quad A_{13} = \begin{bmatrix} 5 & 3 \\ 6 & 7 \\ 2 & 7 \end{bmatrix},$$

$$A_{21} = \begin{bmatrix} 7 & 6 & 5 \\ 7 & 8 & 7 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 4 & 3 \\ 8 & 9 \end{bmatrix}, \quad A_{23} = \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}.$$

Let $A, B \in K^{m \times n}$, and let

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rs} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & \cdots & B_{1s} \\ \vdots & & \vdots \\ B_{r1} & \cdots & B_{rs} \end{bmatrix},$$

$A_{ij}, B_{ij} \in K^{m_i \times n_j}$, $m_1 + \cdots + m_r = m$, $n_1 + \cdots + n_s = n$. In this case we say that the block matrices A and B are of the same type. Clearly the sum $A + B$ is obtainable by adding submatrices blockwise:

$$A + B = \begin{bmatrix} A_{11} + B_{11} & \cdots & A_{1s} + B_{1s} \\ \vdots & & \vdots \\ A_{r1} + B_{r1} & \cdots & A_{rs} + B_{rs} \end{bmatrix}.$$

In the case of multiplication of two block matrices, the partition should be **conformable**, that is, it is such that the multiplication of submatrices make sense. If $A \in K^{m \times q}$ is a block matrix:

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rs} \end{bmatrix},$$

with $A_{ij} \in K^{m_i \times q_j}$, $m_1 + \cdots + m_r = m$, $q_1 + \cdots + q_s = q$. If $B \in K^{q \times n}$ then the conformable partition of B with respect to A for AB is

$$B = \begin{bmatrix} B_{11} & \cdots & B_{1t} \\ \vdots & & \vdots \\ B_{s1} & \cdots & B_{st} \end{bmatrix},$$

where $B_{ij} \in K^{q_i \times n_j}$, $n_1 + \cdots + n_t = n$. In this case we obtain

$$AB = \begin{bmatrix} C_{11} & \cdots & C_{1t} \\ \vdots & & \vdots \\ C_{r1} & \cdots & C_{rt} \end{bmatrix},$$

where $C_{ij} = \sum_{k=1}^{s} A_{ik} B_{kj}$.

A square matrix A is called **upper triangular** (respectively, **lower triangular**) if $[A]_{ij} = 0$ for $i > j$ (respectively, $i < j$). A square matrix

A is called a **diagonal matrix** if it is upper triangular as well as lower triangular, that is, $[A]_{ij} = 0$ for $i \neq j$.

A block matrix

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1r} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rr} \end{bmatrix}$$

with each A_{ii} a square matrix, then:

(i) A is **block diagonal** if $A_{ij} = 0$ for $i \neq j$;
(ii) A is **block upper triangular** if $A_{ij} = 0$ for $i > j$;
(iii) A is **block lower triangular** if $A_{ij} = 0$ for $i < j$.

If A is an $n \times n$ diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$, then we write $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Similarly, $A = \operatorname{diag}(A_1, \ldots, A_r)$, where each A_i is a $n_i \times n_i$ matrix, denotes a block diagonal matrix with diagonal entries $A_1, \ldots, A_r$.

The $n \times n$ matrix $I_n = \operatorname{diag}(1, \ldots, 1)$ is the **identity matrix**. If $A \in K^{m \times n}$, then $I_m A = A = A I_n$.

An $n \times n$ matrix A is said to be **invertible** if there exists an $n \times n$ matrix B such that $AB = I_n = BA$. The matrix B is called an **inverse** of A. If A is invertible then it has a unique inverse. Indeed, if B and C are inverses of A, then $B = BI_n = B\,(AC) = (BA)\,C = I_n C = C$. We denote the inverse of A by A^{-1}. Clearly, $\left(A^{-1}\right)^{-1} = A$.

Every square matrix is not invertible. If A is a square matrix whose r–th row is zero, then for any matrix B, the r–th row of AB is also zero, and so A is not invertible. The sum of two invertible matrices need not be an invertible matrix. However, if A and B are invertible matrices, then so is AB, and

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

Thus, a finite product of invertible matrices is also an invertible matrix.

The **transpose** of an $m \times n$ matrix is an $n \times m$ matrix A^t such that $\left[A^t\right]_{ij} = [A]_{ji}$. If $A \in \mathbb{C}^{n \times n}$, then A^* denotes the **conjugate transpose** of A, that is, $[A^*]_{ij} = \left[\overline{A}\right]_{ji}$, where bar denotes the conjugate of a complex number. A square matrix A is **symmetric** (respectively, **Hermitian**) if $A^t = A$ (respectively, $A^* = A$), and **skew–symmetric** (respectively, **skew–Hermitian**) if $A^t = -A$ (respectively, $A^* = -A$).

The following proposition gives elementary properties of transpose.

Proposition 1.1. *Let K be a field. If $A, B \in K^{m \times n}$ and $\lambda \in K$, then*

(i) $\left(A^t\right)^t = A$;
(ii) $(A + B)^t = A^t + B^t$;

(iii) $(\lambda A)^t = \lambda A^t$.

If $A \in K^{m\times q}$ *and* $B \in K^{q\times n}$, *then*

(iv) $(AB)^t = B^t A^t$.

PROOF. (i) $\left[\left(A^t\right)^t\right]_{ij} = \left[A^t\right]_{ji} = [A]_{ij}$ for all $i = 1, \ldots, m,\ j = 1, \ldots, n$. Hence $\left(A^t\right)^t = A$. (ii) and (iii) are easy.

(iv) $\left[(AB)^t\right]_{ij} = [AB]_{ji} = \sum_{k=1}^q [A]_{jk} [B]_{ki} = \sum_{k=1}^q \left[B^t\right]_{ik} \left[A^t\right]_{kj} = \left[B^t A^t\right]_{ij}$. Hence $(AB)^t = B^t A^t$. □

Similar statement hold for conjugate transpose.

PROPOSITION 1.2. *If* $A, B \in \mathbb{C}^{m\times n}$, *then for* $\lambda \in \mathbb{C}$*:*

(i) $(A^*)^* = A$;
(ii) $(A+B)^* = A^* + B^*$;
(iii) $(\lambda A)^* = \overline{\lambda} A^*$.

If $A \in \mathbb{C}^{m\times q}$ *and* $B \in \mathbb{C}^{q\times n}$, *then*

(iv) $(AB)^* = B^* A^*$.

Let $E(m,n)_{ij}$ be an $m \times n$ matrix whose (i,j)–th entry is 1 and all other entries zero, and let $e(n)_i$ be a column vector of length n with 1 at i–th row and all other entries zero.

PROPOSITION 1.3. *Let* $A \in K^{m\times n}$.

(i) $Ae(n)_j$ *is the* j*–th column of* A*;*
(ii) $e(m)_i^t A$ *is the* i*–th row of* A*;*
(iii) $e(m)_i^t Ae(n)_j = [A]_{ij}$*;*
(iv) $e(n)_i^t e(n)_j = \delta_{ij}$ *(Kronecker delta);*
(v) $e(m)_i e(n)_j^t = E(m,n)_{ij}$.

PROOF. Exercise. □

We normally do not write labels m and n for $E(m,n)_{ij}$ and $e(n)_i$ as the size of these matrices is generally clear from the context. In $K^{m\times n}$ the mn matrices E_{ij} $(i = 1, \ldots, m,\ j = 1, \ldots, n)$ are called the **matrix units**. A matrix $A \in K^{m\times n}$ can be uniquely expressed as a linear combination of these mn matrix units:

$$A = \sum_{j=1}^{n} \sum_{i=1}^{m} [A]_{ij} E_{ij}.$$

In $K^{n\times n}$,

$$E_{ij} E_{rs} = \delta_{jr} E_{is}.$$

This identity is a consequence of Proposition 1.3 and the associativity of matrix multiplication: $E_{ij}E_{rs} = \left(e_i e_j^t\right)\left(e_r e_s^t\right) = e_i\left(e_j^t e_r\right)e_s^t = \delta_{jr} e_i e_s^t = \delta_{jr} E_{is}$.

If A is a square matrix, then the **trace** of A, denoted by $\operatorname{tr} A$, is the sum of diagonal entries of A. If $A \in K^{n\times n}$, then $\operatorname{tr} A \in K$. Thus

$$\operatorname{tr} A = \sum_{i=1}^{n} [A]_{ii}\,.$$

PROPOSITION 1.4. *Let K be a field, let $\lambda \in K$ and $A, B, P \in K^{n\times n}$, where P is invertible. Then:*

(i) $tr(A+B) = tr(A) + tr(B)$;
(ii) $tr(\lambda A) = \lambda\, tr(A)$;
(iii) $tr(AB) = tr(BA)$;
(iv) $tr\left(P^{-1}AP\right) = tr(A)$.

PROOF. (i) and (ii) are easy.
(iii)

$$\begin{aligned}\operatorname{tr}(AB) &= \sum_{i=1}^{n} [AB]_{ii} = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} [A]_{ij}[B]_{ji}\right)\\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} [B]_{ji}[A]_{ij}\right) = \sum_{j=1}^{n} [BA]_{jj} = \operatorname{tr}(BA).\end{aligned}$$

(iv) By (iii), $\operatorname{tr}\left(P^{-1}AP\right) = \operatorname{tr}\left((AP)P^{-1}\right) = \operatorname{tr}(A)$. □

2. Elementary operations on matrices

The following operations on rows (respectively, columns) of a matrix are called **elementary row** (respectively, **column**) **operations**:

I interchange of two rows (respectively, columns);
II multiplication of a row (respectively, column) by a nonzero scalar;
III adding to a row (respectively, column) a scalar multiple of another row (respectively, column).

We first show that elementary row (respectively, column) operations on an $m \times n$ matrix A can be performed by premultiplying (respectively, postmultiplying) A by a suitable $m \times m$ (respectively, $n \times n$) matrix.

Observe that if E_{rs} is an $m \times m$ matrix unit, then by Proposition 1.3, $E_{rs} = e_r e_s^t$, and the k–th row of $E_{rs}A$ is $e_k^t\left(E_{rs}A\right) = \left(e_k^t e_r\right)e_s^t A = \delta_{kr} e_s^t A$. Thus, $E_{rs}A$ is an $m \times n$ matrix whose r–th row is identical to the s–th row of A, and all other entries zero.

$$E_{rs}A = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ [A]_{s1} & \cdots & [A]_{sn} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \rightarrow r\text{-th row.} \\ \\ \\ \\ \end{matrix}$$

Similarly, if E_{rs} is an $n \times n$ matrix unit, then the k–th column of AE_{rs} is $(AE_{rs})\, e_k = Ae_r \left(e_s^t e_k\right) = \delta_{sk} Ae_r$. Thus, AE_{rs} has s–th column identical to the r–th column of A, and all other entries zero.

$$AE_{rs} = \begin{bmatrix} 0 & \cdots & 0 & [A]_{1r} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & [A]_{mr} & 0 & \cdots & 0 \end{bmatrix}.$$
$$\downarrow$$
$$s\text{-th column}$$

Define the following $n \times n$ square matrices:

1. $\mathcal{E}(r,s)$, the matrix obtained by interchanging the r–th and s–th row of the identity matrix, called an **elementary permutation matrix**. $\mathcal{E}(r,s) = I_n + E_{rs} + E_{sr} - E_{rr} - E_{ss}$;
2. $\mathcal{E}_r(\lambda) = I_n + (\lambda - 1)\, E_{rr}$, $\lambda \neq 0$, called an **elementary dilation**;
3. $\mathcal{E}_{rs}(\lambda) = I_n + \lambda E_{rs}$, $r \neq s$, called an **elementary transvection**.

We call the above matrices as **elementary matrices**. Elementary matrices are invertible; and the inverse of an elementary matrix is an elementary matrix of the same type: $\mathcal{E}(r,s)^{-1} = \mathcal{E}(r,s)$, $\mathcal{E}_r(\lambda)^{-1} = \mathcal{E}_r\left(\lambda^{-1}\right)$, $\mathcal{E}_{rs}(\lambda)^{-1} = \mathcal{E}_{rs}(-\lambda)$. A matrix which is a product of elementary permutation matrices is called a **permutation matrix**. If P is a permutation matrix, then $P^{-1} = P^t$. Thus, using the above observations, we can prove the next result.

PROPOSITION 2.1. *Let A be an $m \times n$ matrix.*

(i) $\mathcal{E}(r,s)\, A$ *(respectively, $A\mathcal{E}(r,s)$) is an $m \times n$ matrix which is same as A except that the r–th and s–th rows (respectively, columns) are interchanged, that is, it is obtained from A by elementary row (respectively, column) operation of type I.*

(ii) $\mathcal{E}_r(\lambda)\, A$ *(respectively, $A\mathcal{E}_r(\lambda)$) is an $m \times n$ matrix which is same as A except that the r–th row (respectively, column) is multiplied by λ, that* *is, it is obtained from A by applying elementary row (respectively, column) operation of type II.*

(iii) $\mathcal{E}_{rs}(\lambda)A$ *(respectively,* $A\mathcal{E}_{rs}(\lambda)$*) is an* $m \times n$ *matrix which is same as* A *except that the* r*–th row (respectively,* s*–th column) is replaced by* r*–th row (respectively,* s*–th column) of* A *plus* λ *times the* s*–th row (respectively,* r*–th column), that is, it is obtained from* A *by applying elementary row (respectively, column) operation of type III.*

Obviously, we are premultiplying A by elementary matrices of size $m \times m$, and postmultiplying A by elementary matrices of size $n \times n$.

An $m \times n$ matrix is said to be a **row reduced echelon matrix** if:

(i) all zero rows are in the bottom position;
(ii) the leading entry, that is, the first nonzero entry, of each nonzero row is 1.
(iii) if the first t rows are nonzero, and the leading entry of each i–th row is at the k_i–th column, for $i = 1, \ldots, t$, then all other entries of k_i–th column are zero, and $k_i > k_{i-1}$, $i = 2, \ldots, t$.

The zero matrix is deemed to be a row reduced echelon matrix. Examples of row reduced echelon matrices are

$$\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 6 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 5 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and the identity matrix I_n.

Matrices A and B of the same size are said to be **row equivalent** if one can be obtained from the other by elementary row operations. Thus, A and B are row equivalent if and only if there are elementary matrices $P_1, \ldots, P_k$ such that $B = P_k \cdots P_1 A$.

THEOREM 2.2. *An* $m \times n$ *matrix* A *is row equivalent to a row reduced echelon matrix of size* $m \times n$.

PROOF. Clearly, we can assume that A is a nonzero matrix. Among all nonzero rows of A select a row whose leading entry is at the left most column j_1. Thus, all the columns of A which are on the left of column j_1 (if any) are zero. Interchange this row with the first row of A (if required), and we have the matrix A_1 of the following form:

$$A_1 = \begin{bmatrix} 0 & \cdots & 0 & a_1 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_m & * & \cdots & * \end{bmatrix}, \quad a_1 \neq 0.$$

$\downarrow$

j_1-th column

$*$ denotes the entries of A_1 whose precise values are of no concern.

Multiply the first row by a_1^{-1} so that the leading entry of the first row is 1. Then apply elementary row operations of type III to get A_2:

$$A_2 = \begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{bmatrix}.$$

$\downarrow$

j_1-th column

If there are no nonzero rows below the first row of A_2 then it is a row reduced echelon matrix. Otherwise, among all nonzero rows of A_2 which are below the first row select a row whose leading entry is at the left most column j_2. Thus, all columns of A_2 which are on the left of j_2-th column have all their entries zero except in the first row. Interchange this row with the second row of A_2, and we have:

$$A_3 = \begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & b_1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & b_2 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & b_m & * & \cdots & * \end{bmatrix}, \quad b_2 \neq 0.$$

$\downarrow$ $\downarrow$

j_1-th column j_2-th column

Next, multiply the second row by b_2^{-1} so that the leading entry of the second row is 1. Then apply elementary row operations of type III to get:

$$A_4 = \begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & * & \cdots & * \end{bmatrix}.$$

$$\downarrow \qquad\qquad \downarrow$$

j_1-th column $\qquad j_2$-th column

If all the rows below the second row of A_4 are zero then A_4 is a row reduced echelon matrix. Otherwise, we continue this procedure to get to reduce A to a row reduced echelon matrix. □

If $A \in K^{m\times n}$ is row equivalent to a row reduced echelon matrix E then E is called **row reduced echelon form** of A. Next theorem proves that row reduced echelon form for a given matrix is unique.

THEOREM 2.3. *Row reduced echelon form for a matrix is unique.*

PROOF. Let $A \in K^{m\times n}$ and let R and S be row reduced echelon forms of A. Then there is an invertible $m \times m$ matrix M such that $MR = S$.

Let l-th column be the first column in R and S which is not the same, that is, $Re_j = Se_j$, for $j = 1, 2, \ldots, l-1$ and $Re_l \neq Se_l$. Let R' be a matrix such that it has all the columns with leading entry which appear in R before the l-th column of R. Thus if there are q leading entries before the l-th column of R, then $R' = \begin{bmatrix} e_1 & \cdots & e_q & Re_l \end{bmatrix}$, an $m \times (q+1)$ matrix whose first q columns are $e_1, \ldots, e_q$ and $(q+1)$-th column is the l-th column of R. Since these are also the columns of S appearing exactly at the same place, write $S' = \begin{bmatrix} e_1 & \cdots & e_q & Se_l \end{bmatrix}$. The matrices R' and S' are row reduced echelon forms and $MR' = S'$. Therefore $Me_i = e_i$ for $i = 1, \ldots, q$.

It follows by the definition of the row reduced echelon form that Re_l and Se_l are either both of the form $\begin{bmatrix} a \\ 0 \end{bmatrix}$, $a \in K^q$; or one of this type and the other e_{q+1}.

If $Re_l = \begin{bmatrix} c \\ 0 \end{bmatrix}$, $c \in K^q$ and $Se_l = e_{q+1}$, then equating the $(q+1)$-th columns of MR' and S':

$$e_{q+1} = S'e_{q+1} = MR'e_{q+1} = M\begin{bmatrix} c \\ 0 \end{bmatrix} = \begin{bmatrix} c \\ 0 \end{bmatrix},$$

a contradiction. Similar proof when $Re_l = e_{q+1}$. Finally, if $Re_l = \begin{bmatrix} c \\ 0 \end{bmatrix}$ and $Se_l = \begin{bmatrix} c' \\ 0 \end{bmatrix}$, $c, c' \in K^q$, then $\begin{bmatrix} c \\ 0 \end{bmatrix} = MR'e_{q+1} = S'e_{q+1} = \begin{bmatrix} c' \\ 0 \end{bmatrix}$, and so $c = c'$. Hence, R and S are identical matrices. □

EXAMPLE 2.1. *We illustrate the above procedure by reducing*

$$A = \begin{bmatrix} 0 & 2 & 2 & -1 & 3 \\ 1 & 2 & 5 & 0 & 7 \\ 1 & 4 & 7 & 0 & 13 \\ 3 & -2 & 7 & 2 & 3 \end{bmatrix}$$

to row reduced echelon matrix.

Since, there is a nonzero entry at the first column and the second row, we interchange the first and second rows:

$$A_1 = \mathcal{E}(1,2)A = \begin{bmatrix} 1 & 2 & 5 & 0 & 7 \\ 0 & 2 & 2 & -1 & 3 \\ 1 & 4 & 7 & 0 & 13 \\ 3 & -2 & 7 & 2 & 3 \end{bmatrix}.$$

The leading entry of the first row is already 1. Now we convert all other entries of the first column to zero. We multiply the first row by -3 and add to the fourth row, and multiply the first row by -1 and add to the third row. Thus, we have:

$$A_2 = \mathcal{E}_{31}(-1)\mathcal{E}_{41}(-3)A_1 = \begin{bmatrix} 1 & 2 & 5 & 0 & 7 \\ 0 & 2 & 2 & -1 & 3 \\ 0 & 2 & 2 & 0 & 6 \\ 0 & -8 & -8 & 2 & -18 \end{bmatrix}.$$

For the next step, we can select any of the rows below the first row. Since all elements of the third row are multiple of 2, the leading entry, computations are easier with this row. Therefore, we prefer to interchange the second and the third rows:

$$A_3 = \mathcal{E}(2,3)A_2 = \begin{bmatrix} 1 & 2 & 5 & 0 & 7 \\ 0 & 2 & 2 & 0 & 6 \\ 0 & 2 & 2 & -1 & 3 \\ 0 & -8 & -8 & 2 & -18 \end{bmatrix}.$$