Chapter 2
Matrices and Linear Equations

Matrices play a pivotal role in mathematics, and in turn, in all branches of science, social science, and engineering. This chapter is devoted to the interplay between matrices and systems of linear equations.

2.1 Matrices and Their Algebra

By definition, a $m \times n$ matrix A with entries in a field F is an arrangement

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

of m rows and n columns of elements of F. In short A is denoted by $[a_{ij}]$, where a_{ij} is the entry in the ith row and jth column of A. The ith row

$$(a_{i1}, a_{i2}, \ldots, a_{in})$$

of the matrix A is a vector in F^n, called the ith row vector of A, and it will be denoted by $R_i(A)$. Thus, the matrix A can also be expressed as a column

R. Lal, *Algebra 2*, Infosys Science Foundation Series in Mathematical Sciences,
DOI 10.1007/978-981-10-4256-0_2

$$\begin{bmatrix} R_1(A) \\ R_2(A) \\ \cdot \\ \cdot \\ \cdot \\ R_m(A) \end{bmatrix}$$

of m rows with entries in F^n.
Similarly, if we treat the members of F^m as column vectors, then the jth column

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \cdot \\ \cdot \\ \cdot \\ a_{mj} \end{bmatrix}$$

of the matrix A is a column vector in F^m, called the jth column vector of A, and it will be denoted by $C_j(A)$. As such, the matrix A can also be expressed as a row

$$A = [C_1(A),\ C_2(A),\ \ldots,\ C_m(A)].$$

Thus,

$$A = \begin{bmatrix} 2 & 0 & 1+i & 1-i & 3 \\ 4 & 1+2i & 0 & 1 & i \\ 0 & 8 & 1 & 2 & i \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

is a 4×5 matrix with entries in the field $\mathbb{C}$ of complex numbers.

A matrix A is called a square matrix if the number of rows and columns are same. The matrix

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 4 & 1 & 0 \\ 0 & 8 & 1 \end{bmatrix}$$

is a square 3×3 matrix with entries in the field $\mathbb{R}$ of real numbers.
The set of all $m \times n$ matrices with entries in a field F is denoted by $M_{mn}(F)$. The set of all square $n \times n$ matrices is denoted by $M_n(F)$. We have a binary operation $+$ on $M_{mn}(F)$, called the matrix addition, and which is defined by

$$[a_{ij}] + [b_{ij}] = [c_{ij}],$$

where $c_{ij} = a_{ij} + b_{ij}$.

For example,

$$\begin{bmatrix} 2 & 0 & 1 \\ 4 & 1 & 0 \\ 0 & 8 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 2 \\ 3 & 1 & 0 \\ 5 & 8 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 7 & 2 & 0 \\ 5 & 16 & 2 \end{bmatrix}$$

The $m \times n$ matrix $0_{m\times n}$ all of whose entries are 0 is called the **zero** $m \times n$ matrix. Clearly, the matrix $0_{m\times n}$ is described by the property that for any $m \times n$ matrix A, $A + 0_{m\times n} = A = 0_{m\times n} + A$. Further, if $A = [a_{ij}]$ is a $m \times n$ matrix, then the matrix $-A = [-a_{ij}]$ all of whose entries are the negatives of the corresponding entries of A is called the negative of A, and it is described by the property that $A + (-A) = 0_{m\times n} = (-A) + A$. The proof of the following proposition is an immediate consequence of the corresponding properties of the addition + in F.

Proposition 2.1.1 *The set $M_{mn}(F)$ of $m \times n$ matrices with entries in F is an abelian group with respect to the matrix addition in the sense that it satisfies the following properties:*
(i) The matrix addition + is associative in the sense that
$(A + B) + C = A + (B + C)$
for all A, B and C in $M_{mn}(F)$.
(ii) The matrix addition + is commutative in the sense that
$(A + B) = (B + A)$
for all A, B in $M_{mn}(F)$.
(iii) There is a unique matrix $0_{m\times n}$ in $M_{mn}(F)$ such that $A + 0_{m\times n} = A = 0_{m\times n} + A$ for all A in $M_{mn}(F)$.
(iv) For each matrix A in $M_{mn}(F)$, there is a unique matrix $-A$ in $M_{mn}(F)$ such that $A + (-A) = 0_{mn} = (-A) + A$. ♯

We have an external multiplication $\cdot$ on $M_{mn}(F)$ by scalars in F defined by $a \cdot [a_{ij}] = [b_{ij}]$, where $b_{ij} = a \cdot a_{ij}$. Thus, for example,

$$2 \cdot \begin{bmatrix} 2 & 0 & 1 \\ 4 & 1 & 0 \\ 0 & 8 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 2 \\ 8 & 2 & 0 \\ 0 & 16 & 2 \end{bmatrix}$$

It can be further observed that the triple $(M_{mn}(F), +, \cdot)$ is a vector space over F. Indeed, $(M_{mn}(F), +, \cdot)$ can be identified with the triple $(F^{mn}, +, \cdot)$ under the correspondence $A \longleftrightarrow (R_1(A), R_2(A), \ldots, R_m(A))$ which respects all the operations. Let e_{ij} denote the matrix in which ith row jth column entry is 1 and the rest of the entries are 0. For example, the 3×3 matrix e_{23} is given by

$$e_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

It follows that the set $\{e_{ij} \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}$ corresponds to the standard basis of F^{mn} under the above correspondence. Clearly,

$$[a_{ij}] = \Sigma_{i,j} a_{ij} e_{ij},$$

and

$$\Sigma_{i,j} a_{ij} e_{ij} = 0_{mn}$$

if and only if $a_{ij} = 0$ *for all* i, j. Thus, $\{e_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ is a basis, called the standard basis, of the vector space $M_{mn}(F)$. Thus, the dimension of $M_{mn}(F)$ is $m \cdot n$. In particular, $M_n(F)$ is of dimension n^2.

Apart from the above operations, we have an external operation $\cdot$ from $M_{mn}(F) \times M_{np}(F)$ to $M_{mp}(F)$, called the **matrix multiplication**, defined as follows: Let $A = [a_{ij}]$, $1 \leq i \leq m$, $1 \leq j \leq n$, and $B = [b_{jk}]$, $1 \leq j \leq n$, $1 \leq k \leq p$. Then $A \cdot B = [c_{ik}]$, where $c_{ik} = \Sigma_j a_{ij} b_{jk}$. Thus, for example,

$$\begin{bmatrix} 2 & 0 & 1 \\ 4 & 1 & 0 \\ 0 & 8 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 2 \\ 3 & 1 & 0 \\ 5 & 8 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 10 & 5 \\ 3 & 5 & 8 \\ 29 & 16 & 1 \end{bmatrix}$$

It can be observed easily that the matrix multiplication is distributive over addition from left as well as from right in the sense that $(A + B) \cdot C = A \cdot C + B \cdot C$ and $A \cdot (B + C) = A \cdot B + A \cdot C$. Evidently, $A \cdot 0_{n \times p} = 0_{m \times p}$, and $0_{p \times m} \cdot A = 0_{p \times n}$. Again, since $\Sigma_k(\Sigma_j a_{ij} b_{jk}) c_{kl} = \Sigma_j a_{ij}(\Sigma_k b_{jk} c_{kl})$, it follows that the matrix multiplication is associative in the sense that $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ whenever the products are defined. In particular, we have a multiplication $\cdot$ in $M_n(F)$. Note that matrix multiplication is not commutative, for example,

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

where as

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus, the set $M_n(F)$ of $n \times n$ matrices with entries in F together with matrix addition $+$, the multiplication by scalars, and the matrix multiplication $\cdot$ is an algebra in the sense of the following definition.

Definition 2.1.2 A vector space V over a field F together with an internal multiplication $\cdot$ on V is called an **algebra** over F if the following conditions hold:

1. The internal multiplication $\cdot$ is associative, i.e., $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ for all $x, y, z \in V$.
2. $\cdot$ distributes over addition $+$, i.e., $(x + y) \cdot z = x \cdot z + y \cdot z$, and also $x \cdot (y + z) = x \cdot y + x \cdot z$ for all $x, y, z \in V$.
3. $\alpha(x \cdot y) = (\alpha x) \cdot y = x \cdot (\alpha y)$ for all $\alpha \in F$, and $x, y \in V$.

Let A be a $n \times m$ matrix. The $m \times n$ matrix A^t obtained by interchanging rows and columns of A is called the **transpose** of A. More precisely, if $A = [a_{ij}]$ is a $n \times m$

matrix, then the $m \times n$ matrix $A^t = [b_{ji}]$, where $b_{ji} = a_{ij}$ is called the **transpose** of A. Let $A = [a_{ij}]$ be a $n \times m$ matrix with entries in the field $\mathbb{C}$ of complex numbers. The matrix $\overline{A} = [b_{ij}]$, where $b_{ij} = \overline{a_{ij}}$ (the complex conjugate of a_{ij}) is called the **conjugate** of the matrix A. The matrix $A^\star = \overline{A}^t$ is called the **tranjugate**, also called the **hermitian conjugate** of A

Thus, for example

$$\begin{bmatrix} 2 & 0 & 1 \\ 4 & 1 & 0 \\ 0 & 8 & 1 \end{bmatrix}^t = \begin{bmatrix} 2 & 4 & 0 \\ 0 & 1 & 8 \\ 1 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2+i & i & 1+i \\ 4+i & i & 0 \\ 1-i & 8 & 1+i \end{bmatrix}^\star = \begin{bmatrix} 2-i & 4-i & 1+i \\ -i & -i & 8 \\ 1-i & 0 & 1-i \end{bmatrix}$$

Proposition 2.1.3 *Let A, B be matrices with entries in a field F. Then*

(i) $(A + B)^t = A^t + B^t$
(ii) $(A^t)^t = A$.
(iii) $(a \cdot A)^t = a \cdot A^t$
(iv) $(A \cdot B)^t = B^t \cdot A^t$

provided the relevant sums and the products are defined.
Further, if A, B are matrices with entries in the field $\mathbb{C}$ of complex numbers, then

(v) $(A + B)^\star = A^\star + B^\star$
(vi) $(A^\star)^\star = A$.
(vii) $(a \cdot A)^\star = \overline{a} \cdot A^\star$
(viii) $(A \cdot B)^\star = B^\star \cdot A^\star$

provided the relevant sums and the products are defined.

Proof The identities (i), (ii), and (iii) are evident from the definition. We prove the (iv). Suppose that $A = [a_{ij}]$ is a $n \times m$ matrix, and $B = [b_{jk}]$ is a $m \times p$ matrix. Then, by the definition, $A \cdot B = [c_{ik}]$, where $c_{ik} = \Sigma_j a_{ij} b_{jk} = \Sigma_j v_{kj} u_{ji}$ where $v_{kj} = b_{jk}$ and $u_{ji} = a_{ij}$. By the definition $B^t = [v_{kj}]$, $A^t = [u_{ji}]$, and $(A \cdot B)^t = [w_{ki}]$, where $w_{ki} = c_{ik}$. This shows that the k_{th} row j_{th} column entry of both sides are same. This proves the result. The proofs of the rest of the identities are similar. ♯

2.2 Types of Matrices

1. **Identity matrix**. The $n \times n$ matrix all of whose diagonal entries are 1 and off diagonal entries are 0 is called the **identity** matrix of order n, and it is denoted by I_n. For example,

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

It can be checked that $I_n \cdot A = A = A \cdot I_m$ for every $n \times m$ matrix A. Indeed, if C is a $n \times m$ matrix such that $C \cdot A = A$ for every $n \times m$ matrix A, then $C = I_n$.
2. **Diagonal matrix**. A matrix $A = [a_{ij}]$ is called a **diagonal matrix** if all off diagonal entries are 0. Thus, $[a_{ij}]$ is a diagonal matrix if $a_{ij} = 0$ for all $i \neq j$. The diagonal matrix whose ith row ith column entry is α_i is denoted by $Diag(\alpha_1, \alpha_2, \ldots, \alpha_n)$. For example,

$$Diag(1,\ 2,\ 3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

The effect of multiplying the diagonal matrix $diag(\alpha_1, \alpha_2, \ldots, \alpha_n)$ to a $n \times m$ matrix A from left is to multiply the ith row by α_i. Thus $diag(\alpha_1, \alpha_2, \ldots, \alpha_n) \cdot [a_{ij}] = [b_{ij}]$, where $b_{ij} = \alpha_i a_{ij}$. Similarly, the effect of multiplying this matrix to a $m \times n$ matrix A from right is the same as multiplying the ith column by α_i. In particular, $diag(\alpha_1, \alpha_2, \ldots, \alpha_n) \cdot diag(\beta_1, \beta_2, \ldots \beta_n) = diag(\alpha_1\beta_1, \alpha_2\beta_2, \ldots, \alpha_n\beta_n)$.
3. **Scalar matrix**. A $n \times n$ diagonal matrix all of whose diagonal entries are same is called a **scalar** matrix. Thus, a scalar matrix is of the form αI_n, and effect of multiplying this matrix to a matrix A is αA.
4. **Symmetric matrix**. A matrix A is called a **symmetric matrix** if $A^t = A$. Thus, a diagonal matrix is a symmetric matrix. The matrix

$$\begin{bmatrix} 1 & 3 & 2 \\ 3 & 2 & 0 \\ 2 & 0 & 3 \end{bmatrix}$$

is a symmetric matrix. It follows from the Proposition 2.1.3 that sum of two symmetric matrices are symmetric, scalar multiple of a symmetric matrix is a symmetric matrix. Thus, the set $S_n(F)$ of all $n \times n$ symmetric matrices forms a subspace of $M_n(F)$. For all matrices A, AA^t is a symmetric matrix. For a square matrix A, $A + A^t$ is a symmetric matrix. Product of two symmetric matrices is symmetric if and only if they commute.
5. **Skew symmetric matrix**. A matrix A is called a **skew symmetric matrix** if $A^t = -A$. For example, the matrix

$$\begin{bmatrix} 0 & 3 & 2 \\ -3 & 0 & 0 \\ -2 & 0 & 0 \end{bmatrix}$$

is a skew symmetric matrix. It follows from the Proposition 2.1.3 that sum of two skew symmetric matrices are skew symmetric, scalar multiple of a skew symmetric matrix is a skew symmetric matrix. Thus, the set $SS_n(F)$ of all $n \times n$ skew symmetric matrices forms a subspace of $M_n(F)$. $A - A^t$ is skew symmetric for all square matrices A. Product of two skew symmetric matrices is skew symmetric if and only if they anti commute in the sense that $A \cdot B = -B \cdot A$. Also observe that the diagonal entries of a skew symmetric matrices are 0.

Every square matrix A with entries in a field F can be uniquely represented as sum $A = \frac{A+A^t}{2} + \frac{A-A^t}{2}$ of a symmetric matrix $\frac{A+A^t}{2}$ and a skew symmetric matrix $\frac{A-A^t}{2}$ (prove the uniqueness of the representation).

6. **Hermitian matrix**. A matrix A with entries in the field $\mathbb{C}$ of complex numbers is called a **hermitian matrix** (also termed as **self adjoint**) if $A^\star = A$. Thus, a matrix A with real entries is Hermitian if and only if it is symmetric. The matrix

$$\begin{bmatrix} 1 & 3+i & 2 \\ 3-i & 2 & i \\ 2 & -i & 3 \end{bmatrix}$$

is a Hermitian matrix. Evidently, all diagonal entries of Hermitian matrices are real. It follows from the Proposition 2.1.3 that sum of two Hermitian matrices are Hermitian. However, only real scalar multiple of a Hermitian matrix is a Hermitian matrix. For all matrices A, $AA^\star$ is a Hermitian matrix. For a square matrix A, $A + A^\star$ is also a Hermitian matrix. Product of two Hermitian matrices is Hermitian if and only if they commute.

7. **Skew-Hermitian matrix**. A matrix A with entries in the field $\mathbb{C}$ of complex numbers is called a **skew-Hermitian matrix** if $A^\star = -A$. Thus, a matrix A with real entries is skew-Hermitian if and only if it is skew symmetric. The matrix

$$\begin{bmatrix} i & 3i-1 & 2 \\ 3i+1 & 2i & -1 \\ 2 & 1 & 3i \end{bmatrix}$$

is a skew-Hermitian matrix. Evidently, all diagonal entries of skew-Hermitian matrices are purely imaginary. It follows from the Proposition 2.1.3 that sums of two skew-Hermitian matrices are skew-Hermitian. However, only real scalar multiple of a skew-Hermitian matrix is a skew-Hermitian matrix. Observe that a matrix A is skew-Hermitian if and only if iA is a Hermitian matrix. For all matrices A, $iAA^\star$ is a skew-Hermitian matrix. For a square matrix A, $A - A^\star$ is also a skew-Hermitian matrix. Product of two skew-Hermitian matrices is skew-Hermitian if and only if they anticommute in the sense that $AB = -BA$.

Every square matrix A with entries in the field $\mathbb{C}$ of complex numbers can be uniquely represented as sum $A = \frac{A+A^\star}{2} + \frac{A-A^\star}{2}$ of a Hermitian matrix $\frac{A+A^\star}{2}$, and a skew-Hermitian matrix $\frac{A-A^\star}{2}$ (prove the uniqueness of the representation). In turn, it follows that every square matrix A with entries in the field $\mathbb{C}$ of complex numbers can be uniquely represented as $A = B + iC$, where B and C are Hermitian matrices.

8. **Nonsingular matrices**. A $n \times n$ matrix A is called a **nonsingular matrix** (also called an **invertible matrix**) if there is a $n \times n$ matrix B such that $A \cdot B = I_n = B \cdot A$.

Note that such a B, if exists, will be unique, for if B_1 and B_2 are such matrices, then $B_1 = B_1 \cdot I_n = B_1 \cdot (A \cdot B_2) = (B_1 \cdot A) \cdot B_2 = I_n \cdot B_2 = B_2$. If A is an invertible matrix, then the unique B such that $A \cdot B = I_n = B \cdot A$ is called the **Inverse** of A, and it is denoted by A^{-1}. Following are some simple observations:
(i) The identity matrix I_n is invertible and $I_n^{-1} = I_n$.
(ii) Consider a diagonal matrix $diag(\alpha_1, \alpha_2, \ldots, \alpha_n)$. As already observed in 2, $diag(\alpha_1, \alpha_2, \ldots, \alpha_n) \cdot [a_{ij}] = [b_{ij}]$, where $b_{ij} = \alpha_i a_{ij}$. Thus, $diag(\alpha_1, \alpha_2, \ldots, \alpha_n) \cdot [a_{ij}] = I_n$ if and only if $\alpha_i a_{ij} = 1$ for $j = i$, and 0 other wise. This is so if and only if $\alpha_i \neq 0$, $a_{ii} = \alpha_i^{-1}$ for each i, and $a_{ij} = 0$ for all $i \neq j$. This shows that $Diag(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is invertible if and only if each $\alpha_i \neq 0$, and then its inverse is $Diag(\alpha_1^{-1}, \alpha_2^{-1}, \ldots, \alpha_n^{-1})$.
(iii) Let A and B be invertible $n \times n$ matrices. Then, $(AB)(B^{-1}A^{-1}) = I_n = (B^{-1}A^{-1})(AB)$. This shows that AB is also invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

In due course, we shall describe an algorithms to check if a matrix is invertible, and then to find its inverse.

9. **Triangular matrices**. A square matrix A is said to be an **upper (lower) triangular** matrix if all its below (above) diagonal entries are 0. More precisely, a $n \times n$ matrix $A = [a_{ij}]$ is called an **upper (lower) triangular** matrix if $a_{ij} = 0 \textit{ for all } i > j (i < j)$. It is called **strictly upper (lower)** triangular if in addition to that all the diagonal entries are also 0. For example,

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

is an upper triangular matrix.

Clearly, the sum of any two **upper (lower) triangular** matrices is an **upper (lower) triangular** matrix. Also a scalar multiple of an **upper (lower) triangular** matrix is a **upper (lower) triangular** matrix. Thus, the set $T_+(n, F)(T_-(n, F))$ of upper (lower) triangular matrices forms a subspace of $M_n(F)$.

Further, $T_+(n, F)(T_-(n, F))$ is closed under matrix multiplication: For, let $A = [a_{ij}]$ and $B = [b_{jk}]$ be upper triangular matrices. Then $a_{ij} = 0 = b_{jk} \textit{ for all } i > j > k$. Let $A \cdot B = [c_{ik}]$. Then $c_{ik} = \Sigma_j a_{ij} b_{jk} = 0 \textit{ for all } i > k$.

Next, let $A = [a_{ij}] \in T_+(n, F)$ be a nonsingular matrix. Then there is a matrix $B = [b_{ij}]$ such that $B \cdot A = I_n$. Equating the first row first column entry from both side we get $b_{11}a_{11} = 1$. But then $a_{11} \neq 0$ and $b_{11} = a_{11}^{-1}$. Equating second row first column entry, we obtain that $b_{21}a_{11} = 0$. Hence $b_{21} = 0$. Similarly, equating ith row 1_{st} column entry we obtain that $b_{i1}a_{11} = 0$, and so $b_{i1} = 0 \textit{ for all } i > 1$. Equating the 1_{st} row 2_{nd} column entry, we get that $b_{11}a_{12} + b_{12}a_{22} = 0$, and equating the 2_{nd} row 2_{nd} column entry, we get $b_{22}a_{22} = 1$. Thus $a_{22} \neq 0$, $b_{22} = a_{22}^{-1}$, and $b_{12} = a_{22}^{-1}a_{11}^{-1}a_{12}$. Proceeding in this way we obtain that all the diagonal entries a_{ii} of A are nonzero, and then we can solve b_{ij} to get the inverse of A. We also observe that the inverse of A is also a member of $T_+(n, F)$. For example, consider the upper triangular matrix

$$\begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

all of whose diagonal entries are nonzero. We find its inverse. Suppose that

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then we have the following equations:
$2a_{11} = 1,\ 4a_{11} + 2a_{12} = 0,\ 6a_{11} + 3a_{13} = 0,$
$2a_{21} = 0,\ 4a_{21} + 2a_{22} = 1,\ 6a_{21} + 3a_{23} = 0,$
$2a_{31} = 0,\ 4a_{31} + 2a_{32} = 0,\ 3a_{33} = 1$
Solving, we get that $a_{11} = \frac{1}{2},\ a_{12} = -1 = a_{13},\ a_{21} = a_{31} = a_{32} = 0,\ a_{23} = 0,\ a_{22} = \frac{1}{2},\ a_{33} = \frac{1}{3}$. Thus, the inverse of the said matrix is

$$\begin{bmatrix} \frac{1}{2} & -1 & -1 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$

Block multiplication
We can multiply two matrices by using suitable blocks of their submatrices. More explicitly, let A be a $m \times n$ matrix, and B a $n \times p$ matrix. Suppose that $m = m_1 + m_2 + \cdots + m_r,\ n = n_1 + n_2 + \cdots n_s$, and $p = p_1 + p_2 + \cdots + p_t$, where m_i, n_j and p_k are positive integers. Then A and B can be expressed uniquely as

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{bmatrix},$$

where A_{ij} is a $m_i \times n_j$ matrix and

$$B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1t} \\ B_{21} & B_{22} & \cdots & B_{2t} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ B_{s1} & B_{s2} & \cdots & B_{st} \end{bmatrix},$$

where B_{jk} is $n_j \times p_k$ matrix. Further, then

$$A \cdot B = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1t} \\ C_{12} & C_{22} & \cdots & C_{2t} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ C_{r1} & C_{r2} & \cdots & C_{rt} \end{bmatrix},$$

where $C_{ik} = \Sigma_{j=1}^{s} A_{ij} B_{jk}$.

2.3 System of Linear Equations

A system of m linear equations in n unknowns $x_1, x_2, \ldots, x_n$ over a field F is given by

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = b_2 \\ \cdot \quad \cdot \quad \cdot \quad \cdots\cdots \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdots\cdots \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdots\cdots \quad \cdot \quad \cdot \quad \cdot \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = b_n \end{pmatrix}, \tag{2.1}$$

where $a_{ij} \in F$.

Example 2.3.1 Following is a system of two linear equations in three unknowns over the field of real numbers:

$$3x_1 + 2x_2 + x_3 = 1.$$

$$x_1 + x_2 + x_3 = 2.$$

We say that a *n-tuple* $(a_1, a_2, \ldots, a_n)$ in F^n is a solution of the system (2.1) of linear equations if $x_1 = a_1, x_2 = a_2, \ldots, x_n = a_n$ satisfies all the equations in the system (2.1). Thus, $(-2, 3, 1)$ is a solution of the system of linear equations in the above example. $(-3, 5, 0)$ is also a solution to the above system. Indeed, there are infinitely many solutions which can be parametrized in terms of x_3 as $(x_3 - 3, 5 - 2x_2, x_3)$. Clearly, this represents a line.

Example 2.3.2 The system

$$x_1 + 2x_2 + 3x_3 = 1.$$

$$x_1 + x_2 + 3x_3 = 2.$$

$$4x_1 + 6x_2 + 12x_3 = 5.$$

of linear equations has no solution (why?).

where as

Example 2.3.3 The system

$$x_1 + 2x_2 = 1.$$

$$2x_1 + 2x_2 = a.$$

has a unique solution for all a (why?).

Definition 2.3.4 A system of linear equations is said to be **consistent** if it has a solution. It is said to be **inconsistent** otherwise.

The Example 2.3.1 is consistent having infinitely many solutions, the Example 2.3.2 is inconsistent, whereas Example 2.3.3 is consistent with unique solution.

Most of the problems in real life, in engineering, in industries, in social life, and in medical science can be modeled in terms of systems of linear equations. As such, describing and interpreting the solutions of a system of linear equations is one of the main themes of linear algebra. In the following few sections we shall concentrate on this.

The system (2.1) of m linear equations in n unknowns can be expressed in a single matrix equation

$$A\overline{x}^t = \overline{b}^t \tag{2.2}$$

where $A = [a_{ij}]$ is the $m \times n$ matrix whose i_{th} row j_{th} column entry is a_{ij}, $\overline{x} = [x_1, x_2, \ldots, x_n] \in F^n$ is the $1 \times n$ row matrix of unknowns, and $\overline{b} = [b_1, b_2, \ldots, b_m] \in F^m$ is the $1 \times m$ matrix.

Thus, the system of linear equations in Example 2.3.1 can be expressed as

$$\begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

The matrix A in (2.2) is called the **coefficient matrix** of the system (2.1) of linear equations, and the $m \times (n+1)$ matrix $A^+ = [A\,\overline{b}^t]$ whose first n columns are those of A, and the last $(n+1)_{th}$ column is $\overline{b}^t$, is called the **augmented matrix** of the system of linear equations.

Thus, the coefficient matrix of the Example 2.3.2 is

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 3 \\ 4 & 6 & 12 \end{bmatrix},$$

and the augmented matrix of the example is

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

Definition 2.3.5 A system of linear equations given by the matrix equation

$$A\overline{x}^t = \overline{0}^t \ \ldots . \tag{2.3}$$

is called a **homogeneous system** of linear equations. It is also called the **homogeneous part** of the system of linear equations given by

$$A\overline{x}^t = \overline{b}^t.$$

Proposition 2.3.6 *A homogeneous system of linear equations given by the matrix equation*

$$A\overline{x}^t = \overline{0}^t.$$

is always consistent, and the set of solutions of the homogeneous system is a subspace of F^n.

Proof Let $N(A)$ denote the set of all solutions of $A\overline{x}^t = \overline{0}^t$. Since $A\overline{0}^t = \overline{0}^t$, it follows that $\overline{0} \in N(A)$. Let $\overline{u}, \overline{v} \in N(A)$, and $a, b \in F$. Then $A(a\overline{u} + b\overline{v})^t = aA\overline{u}^t + bA\overline{v}^t = \overline{0}^t$. This shows that $a\overline{u} + b\overline{v} \in N(A)$. It follows that $N(A)$ is a subspace of F^n. ♯

Definition 2.3.7 The subspace $N(A)$ described in the above proposition is called the **solution space** of the system (2.3) of linear equations, and it is also called the **null space** of the matrix A. The dimension of the null space $N(A)$ is called the **nullity** of A, and it is denoted by $n(A)$. If $\{\overline{u_1}, \overline{u_2}, \ldots, \overline{u_{n(A)}}\}$ is a basis of $N(A)$, then any solution of (2.3) is uniquely expressed as $c_1\overline{u_1} + c_2\overline{u_2} + \cdots + c_{n(A)}\overline{u_{n(A)}}$, where $c_1, c_2, \ldots, c_{n(A)}$ are constants in F. As such $c_1\overline{u_1} + c_2\overline{u_2} + \cdots + c_{n(A)}\overline{u_{n(A)}}$ is called a general solution of the homogeneous system (2.3).

A little later, we shall give an algorithm to find $N(A)$, indeed a basis of $N(A)$, and so also a general solution of the system (2.3) of linear equations.

Proposition 2.3.8 *Suppose that the system of linear equations given by the matrix equation*

$$A\overline{x}^t = \overline{b}^t.$$

is consistent, and $\overline{a} = [a_1, a_2, \ldots, a_n]$ is a solution of the above equation. Then the coset $\overline{a} + N(A) = \{\overline{a} + \overline{u} \mid \overline{u} \in N(A)\}$ is the complete set of all solutions of the system of linear equations. In turn, if $\{\overline{u_1}, \overline{u_2}, \ldots, \overline{u_{n(A)}}\}$ is a basis of $N(A)$, then $\overline{a} + c_1\overline{u_1} + c_2\overline{u_2} + \cdots + c_{n(A)}\overline{u_{n(A)}}$ is a general solution of the system of linear equations, where $c_1, c_2, \ldots, c_{n(A)}$ are arbitrary constants.

Proof Since $\overline{a}$ is a solution of $A\overline{x}^t = \overline{b}^t$, $A\overline{a}^t = \overline{b}^t$. If $\overline{u} \in N(A)$, then $A\overline{u}^t = \overline{0}^t$. But, then $A(\overline{a} + \overline{u})^t = (A\overline{a}^t + A\overline{u}^t) = (\overline{b}^t + \overline{0}^t) = \overline{b}^t$. This shows that $\overline{a} + \overline{u}$ is also a solution of $A\overline{x}^t = \overline{b}^t$. Conversely, let $\overline{c}$ be a solution of $A\overline{x}^t = \overline{b}^t$. Then $A\overline{c}^t = \overline{b}^t$. Hence $A(\overline{c} - \overline{a})^t = (A\overline{c}^t - A\overline{a}^t) = \overline{0}$. It follows that $(\overline{c} - \overline{a}) \in N(A)$. This shows that $\overline{c} \in \overline{a} + N(A)$. The rest is an immediate observation. $\sharp$

Definition 2.3.9 The subspace $R(A)$ of F^n generated by the set $\{R_1(A), R_2(A), \ldots, R_m(A)\}$ of the rows of A is called the **row space** of A, and the dimension of $R(A)$ is called the **row rank** of A. Thus, the maximum number of linearly independent rows of a matrix is the row rank of A. Similarly, the subspace $C(A)$ of F^m (the elements of F^m treated as columns) is called the **column space** of A, and the dimension of $C(A)$ is called the **column rank** of A. Again, it follows that the maximum number of linearly independent columns of A is the column rank of A. We shall see, in due course, that row rank is same as column rank, and it is called the **rank**of A. The rank of A is denoted by $r(A)$.

Proposition 2.3.10 *The system of linear equations given by the matrix equation*

$$A\overline{x}^t = \overline{b}^t.$$

is consistent if and only if the column rank of A is same as that of the augmented matrix A^+.

Proof The system of linear equations given by the matrix equation $A\overline{x}^t = \overline{b}^t$ is also expressible as

$$x_1C_1(A) + x_2C_2(A) + \cdots + x_nC_n(A) = \overline{b}^t,$$

where $\overline{x} = [x_1, x_2, \ldots, x_n]$, and $C_i(A)$ denotes the i_{th} column of A. Thus, the equation has a solution if and only if $\overline{b}^t$ is a linear combination of the columns of A. This is equivalent to say that the column space $C(A)$ of A is same as the column space $C(A^+)$ of the augmented matrix A^+. Since $C(A) \subseteq C(A^+)$, this is equivalent to the fact that column rank of A is same as that of A^+. $\sharp$

We shall look at an algorithm to find the rank of a matrix, and also an algorithm to find a general solution of $A\overline{x}^t = \overline{b}^t$ provided it is consistent.

2.4 Gauss Elimination, Elementary Operations, Rank, and Nullity

Definition 2.4.1 Two systems of m linear equations in n unknowns are said to be equivalent if they have same set of solutions.

Example 2.4.2 The system

$$x_1 + 2x_2 = 1$$

$$2x_1 + 2x_2 = a$$

of two linear equations in two unknowns is equivalent to the system

$$x_1 + 2x_2 = 1$$

$$3x_1 + 4x_2 = a + 1,$$

for they have same set of solutions, whereas the system is not equivalent to

$$x_1 + 2x_2 = 1$$

$$2x_1 + 3x_2 = a$$

In what follows, we shall introduce an algorithm called the **Gaussian elimination** to reduce a system of linear equations into an equivalent system of linear equations from which the solution will become apparent.

Definition 2.4.3 Following operations on a system of linear equations are called the **elementary operations** on the system of linear equations, and the corresponding operations on coefficient and augmented matrices are called the **Elementary row operations** on the matrices:

1. Interchange any two equations in the system.
2. Multiply an equation in the system by a nonzero member of the field.
3. Add a nonzero multiple of an equation in the system to another equation in the system.

In turn, the corresponding elementary **row operations** on matrices are:

1. Interchange any two row of the matrix.
2. Multiply a row of the matrix by a nonzero element of the field.
3. Add a nonzero multiple of a row of the matrix to another row.

The following proposition is an immediate observation.

Proposition 2.4.4 *Any two system of linear equations which differ by a finite sequence of elementary operations are equivalent.* ♯

We shall first discuss an algorithm to find the space of solutions of a homogeneous system of linear equations given by the matrix equation $A\overline{x}^t = \overline{0}^t$. More precisely, we derive an algorithm to find a basis of the null space $N(A)$ of A so that every solution of the system is unique linear combination of the basis members.

Proposition 2.4.5 *The null space $N(A)$, and so also the nullity $n(A)$ of a matrix A remain invariant under the elementary row operations.*

Proof Follows from the Proposition 2.4.4. ♯

Proposition 2.4.6 *The row space $R(A)$ and so also the row rank of a matrix A remain invariant under the elementary row operations.*

Proof Interchange of any two rows of a matrix will not change the row space as the set of rows will not change. Since the subspace of F^n generated by the set $\{R_1(A), R_2(A), \ldots, R_m(A)\}$ of rows of A is the same as the subspace of F^n generated by $\{R_1(A), R_2(A), \ldots, aR_j(A), \ldots, R_m(A)\}$ for each nonzero $a \in F$ and $j \leq m$, it follows that the row space of a matrix remains the same if we multiply a row of the matrix by a nonzero member of the field. Finally, since the subspace of F^n generated by the set $\{R_1(A), R_2(A), \ldots, R_m(A)\}$ of rows of A is the same as the subspace of F^n generated by $\{R_1(A), R_2(A), \ldots, R_k(A) + aR_j(A), \ldots, R_m(A)\}$ for each nonzero $a \in F$ and $j \neq k$, it follows that the row space of a matrix remains the same if we add a nonzero multiple of a row to another row. ♯

The column space of a matrix, in general, is not invariant under elementary row operations. However,

Proposition 2.4.7 *The column rank of a matrix remains invariant under elementary row operations.*

Proof Let A be a matrix, and A' a matrix obtained by applying any of the elementary row operations on A. Then evidently,

$$x_1 C_{i_1}(A) + x_2 C_{i_2}(A) + \cdots + x_r C_{i_r}(A) = \overline{0}^t$$

if and only if

$$x_1 C_{i_1}(A') + x_2 C_{i_2}(A') + \cdots + x_r C_{i_r}(A') = \overline{0}^t$$

This means that the maximum number of linearly independent columns of A is same as that of A'. Thus, the column rank of A is same as that of A'. ♯

We shall describe an algorithm to transform a matrix in to a special form, called a **reduced row echelon form**, of the matrix by using elementary row operations, and from which a basis for the null space of the matrix, and also a basis of the row space of the matrix can be easily obtained.

Definition 2.4.8 A $m \times n$ matrix $A = [a_{ij}]$ is said to be a matrix in **reduced row (column) echelon form**, or it is said to be a **reduced row echelon matrix** if the following hold:

(i) The first nonzero entry in each row (column) is 1. This entry is called a **pivot** entry, and the corresponding columns (rows) are called **pivot column (row)** of the matrix. The columns (rows) which are not pivot columns (rows) are called the **free** columns (rows). The unknown variable corresponding to pivot columns are called **pivot variables**, and those corresponding to free columns are called **free Variables**.

(ii) The pivot entry in any row (column) is towards right (bottom) side to the pivot entry in the previous row (column).
(iii) All of the rest of the entries in a pivot column (row) are 0.
(iv) All the zero rows (columns) are towards bottom (right).

Example 2.4.9 The matrix

$$A = \begin{bmatrix} 1\ 2\ 0\ 0\ 2 \\ 0\ 0\ 1\ 0\ 1 \\ 0\ 0\ 0\ 1\ 2 \\ 0\ 0\ 0\ 0\ 0 \end{bmatrix}$$

is in reduced row echelon form. The 1st row 1st column, the 2nd row 3rd column, and the 3rd row 4th column entries are pivot entries, 2nd and 5th columns are free columns. x_1, x_3 and x_4 are pivot variables. x_2 and x_5 are free variables.

Proposition 2.4.10 *Let A be a $m \times n$ matrix with entries in a field F and which is in reduced row echelon form. Suppose that the columns $C_{i_1}(A), C_{i_2}(A), \ldots, C_{i_r}(A)$ with $i_1 < i_2 < \cdots < i_r$ are pivot columns and the columns $C_{j_1}(A), C_{j_2}(A), \ldots, C_{j_s}(A)$ with $j_1 < j_2 < \cdots < j_s$ are free columns. Then,*

(i) the first r rows $R_1(A), R_2(A), \ldots, R_r(A)$ are nonzero rows, and they form a basis of the row space $R(A)$ of A,
(ii) the number of pivots is the row rank of A,
(iii) the pivot columns form a basis of the column space of A,
(iv) row rank of A is the same as the column rank of A. Indeed, it is the number of pivots.

Proof (i) Since each nonzero row contains a unique pivot entry, and the zero rows are towards the bottom, it follows that $R_1(A), R_2(A), \ldots, R_r(A)$ are precisely the nonzero rows of the matrix. Since the pivot entries 1 in $R_1(A), R_2(A), \ldots, R_r(A)$ appear in different columns $i_1, i_2, \ldots, i_r$, it follows that the set $\{R_1(A), R_2(A), \ldots, R_r(A)\}$ of nonzero row of A is linearly independent. As such, it forms a basis of the row space $R(A)$ of A.
(ii) Follows from (i).
(iii) Clearly, the set $\{C_{i_1}(A), C_{i_2}(A), \ldots, C_{i_r}(A)\}$ of pivot columns form a linearly independent set, for the k_{th} row entry in the pivot column $C_{i_k}(A)$ is 1 and the rest of the entries in this column are 0. It is also evident that all the free columns are linear linear combinations of the pivot columns. Indeed,

$$C_{j_l}(A) = a_{1j_l}C_{i_1}(A) + a_{2j_l}C_{i_2}(A) + \cdots + a_{rj_l}C_{i_r}(A).$$

(iv) Follows from (iii). ♯

Proposition 2.4.11 *Consider the homogeneous system of linear equations given by the matrix equation*

$$A\overline{x}^t = \overline{o}^t,$$

where A is a reduced row echelon $m \times n$ matrix with entries in a field F. Suppose that the columns $C_{i_1}(A), C_{i_2}(A), \ldots, C_{i_r}(A)$ with $i_1 < i_2 < \cdots < i_r$ are pivot columns, and the columns $C_{j_1}(A), C_{j_2}(A), \ldots, C_{j_s}(A)$ with $j_1 < j_2 < \cdots < j_s$ are free columns. Then the pivot variables $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$ in the homogeneous system of linear equations are uniquely expressible in terms of free variables $x_{j_1}, x_{j_2}, \ldots, x_{j_s}$ as

$$x_{i_t} = -\sum_{k=1}^{s} a_{tj_k} x_{j_k}.$$

The set $\{\overline{u^1}, \overline{u^2}, \ldots, \overline{u^s}\}$ is a basis for the space $N(A)$ of solutions of the homogeneous system, where $\overline{u^k} = (u_1^k, u_2^k, \ldots, u_n^k)$ is the unique solution of the homogeneous system corresponding to the choice $x_{j_l} = 0$, $l \neq k$, and $x_{j_k} = 1$ of the free variables. Indeed, $u_{j_l}^k = 0$ for $l \neq k$, $u_{j_k}^k = 1$, and $u_{i_t}^k = -a_{tj_k}$. The nullity $n(A) = s$, the number of free variables.

Proof Under the assumption, for all $t \leq r$, $a_{ti_t} = 1$ and $a_{li_t} = 0$ for $l \neq t$. The corresponding homogeneous system of linear equations is given by

$$a_{1i_1}x_{i_1} + a_{1j_1}x_{j_1} + a_{1j_2}x_{j_2} + \cdots + a_{1j_s}x_{j_s} = 0.$$

$$a_{2i_2}x_{i_2} + a_{2j_1}x_{j_1} + a_{2j_2}x_{j_2} + \cdots + a_{2j_s}x_{j_s} = 0.$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a_{ri_1}x_{i_1} + a_{rj_1}x_{j_1} + a_{rj_2}x_{j_2} + \cdots + a_{rj_s}x_{j_s} = 0.$$

the rest of the equations, if any, are the identities

$$0x_1 + 0x_2 + \cdots + 0x_n = 0.$$

Evidently, each pivot variable is uniquely expressible in terms of free variable as described in the proposition. Further, the set $S = \{\overline{u^1}, \overline{u^2}, \ldots, \overline{u^s}\}$ of solutions is a basis of the space $N(A)$ of solutions, for any solution with values $\alpha_1, \alpha_2, \ldots, \alpha_s$ to the free variables $x_{j_1}, x_{j_2}, \ldots, x_{j_s}$ is uniquely expressible as linear combination $\alpha_1\overline{u^1} + \alpha_2\overline{u^2} + \cdots, \alpha_s\overline{u^s}$. The rest is evident. ♯

Proposition 2.4.12 *Consider the system of linear equations given by the matrix equation*

$$A\overline{x}^t = \overline{b}^t,$$

where A is a reduced row echelon $m \times n$ matrix with entries in a field F. Suppose that the columns $C_{i_1}(A), C_{i_2}(A), \ldots, C_{i_r}(A)$ with $i_1 < i_2 < \cdots < i_r$ are pivot columns and the columns

$C_{j_1}(A), C_{j_2}(A), \ldots, C_{j_s}(A)$ with $j_1 < j_2 < \cdots < j_s$ are free columns. Then the system of linear equations is consistent if and only if $b_k = 0$ for all $k \geq r+1$, or equivalently, rank(A) = rank(A^+). Further, then $\overline{v} = (v_1, v_2, \ldots, v_n)$, where $v_{i_t} = -a_{tj_1} + b_t$, $1 \leq t \leq r$, $v_{j_1} = 1$ and $v_{j_l} = 0, 2 \leq l \leq s$, is a particular solution of the above nonhomogeneous system. Finally, a general solution $\overline{x}$ of the system of linear equation is given by

$$\overline{x} = \overline{v} + c_1\overline{u^1} + c_2\overline{u^2} + \cdots + c_s\overline{u^s},$$

where $c_1, c_2, \ldots, c_s$ are arbitrary constants.

Proof From the previous proposition, a general solution of the homogeneous part of the above nonhomogeneous system of linear equations is given by

$$c_1\overline{u^1} + c_2\overline{u^2} + \cdots + c_s\overline{u^s}.$$

Further, the system of linear equations is given by

$$a_{1i_1}x_{i_1} + a_{1j_1}x_{j_1} + a_{1j_2}x_{j_2} + \cdots + a_{1j_s}x_{j_s} = b_1.$$

$$a_{2i_2}x_{i_2} + a_{2j_1}x_{j_1} + a_{2j_2}x_{j_2} + \cdots + a_{2j_s}x_{j_s} = b_2.$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a_{ri_1}x_{i_1} + a_{rj_1}x_{j_1} + a_{rj_2}x_{j_2} + \cdots + a_{rj_s}x_{j_s} = b_r.$$

The rest of the equations, if any, are the identities

$$0x_1 + 0x_2 + \cdots + 0x_n = b_k, \; k \geq r+1.$$

Clearly, the system is inconsistent if $b_k \neq 0$ for any $k \geq r+1$. Now, suppose that $b_k = 0$ *for all* $k \geq r+1$. Putting the free variable $x_{j_1} = 1$, and $x_{j_k} = 0$ for $2 \leq k \leq s$, we get a particular solution $\overline{v} = (v_1, v_2, \ldots, v_n)$, where $v_{i_t} = -a_{tj_1} + b_t$, $1 \leq t \leq r$, $v_{j_1} = 1$, and $v_{j_l} = 0, 2 \leq l \leq s$ of the system. From the Proposition 2.3.8, we get a general solution

$$\overline{x} = \overline{v} + c_1\overline{u^1} + c_2\overline{u^2} + \cdots + c_s\overline{u^s}$$

of the system, where $c_1, c_2, \ldots, c_s$ are arbitrary constants. ♯

Example 2.4.13 Consider the system of linear equations given by the matrix equation

$$A\overline{x}^t = \overline{b}^t,$$

where A is the matrix given by

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The corresponding system of linear equations is given by

$$x_1 + 0x_2 + x_3 + 0x_4 + 2x_5 = b_1.$$

$$0x_1 + x_2 + x_3 + 0x_4 + x_5 = b_2.$$

$$0x_1 + 0x_2 + 0x_3 + x_4 + x_5 = b_3.$$

$$0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = b_4.$$

The matrix A is in reduced row echelon form with the pivot columns C_1, C_2, C_4, and the free columns C_3 and C_5. The nonzero rows R_1, R_2, R_3 form a basis of row space, and the pivot columns C_1, C_2, C_4 of A form a basis of the column space of A. Row rank = 3 = Column rank of A. For the system to be consistent $b_4 = 0$. Assuming that $b_4 = 0$, we find a general solution of the system. We first find a basis of the solution space $N(A)$ of the homogeneous part $A\overline{x}^t = \overline{0}^t$ of the given system of linear equations. x_3 and x_5 are free variables. Putting $x_3 = 1$ and $x_4 = 0$, we get a solution $\overline{u^1} = (-1, -1, 1, 0, 0)$ of the homogeneous part of the system. Further putting $x_3 = 0$ and $x_5 = 1$, we get a solution $\overline{u^2} = (-2, -1, 0, -1, 1)$ of the homogeneous part of the system. The set $\{\overline{u^1}, \overline{u^2}\}$ is a basis of the space $N(A)$ of solutions of the homogeneous part. Nullity of A is 2. Finally, putting $x_3 = 1$ and $x_4 = 0$, we get a particular solution $\overline{v} = (-1 + b_1, -1 + b_2, 1, b_3, 0)$ of the given nonhomogeneous system of linear equations. In turn, a general solution of the given nonhomogeneous system of linear equations is $\overline{v} + c_1\overline{u^1} + c_2\overline{u^2}$.

Observe that a square matrix in reduced row echelon form has no zero rows if and only if all the rows, and so also all columns have pivots, or equivalently, it is the identity matrix. Since a matrix with a zero row is singular, we have the following:

Proposition 2.4.14 *A square matrix in reduced row echelon form is nonsingular if and only if it is the identity matrix.* ♯

Elementary operations on a system of linear equations, or equivalently, elementary row operations on the coefficient and augmented matrices, transform the system into equivalent system of linear equations. Further, if the coefficient matrix of the system of linear equations is in reduced row echelon form, then as observed above, a general solution of the system is easily obtained. As such, it is prompting to discover, if possible, an algorithm to reduce an arbitrary matrix

in to a matrix in reduced row echelon form by using elementary row operations. The following theorem gives an algorithm.

Theorem 2.4.15 *Using elementary row operations, every matrix can be reduced to a matrix in reduced row echelon form.*

Proof Let A be a $m \times n$ matrix. If A is the zero matrix, then it is already in reduced row echelon form. Suppose that A is nonzero matrix. Let j_1 be the least number such that the column $C_{j_1}(A)$ is a nonzero column. Further, let i_1 be the smallest number such that $a_{i_1 j_1} \neq 0$. Interchanging the i_1th row and the first row, we may assume that $a_{1j_1} \neq 0$, and $a_{ik} = 0$ for all $k < j_1$. Multiplying the first row by $a_{1j_1}^{-1}$, we may assume that $a_{1j_1} = 1$, and $a_{ik} = 0$ for all $k < j_1$. Next, adding $-a_{ij_1}$ times the first row to the ith row for each $i \geq 2$, we reduce A to a matrix $[a_{ij}]$, where $a_{1j_1} = 1$, $a_{ij_1} = 0$ for all $i \geq 2$, and $a_{ik} = 0$ for all $k \leq j_1 - 1$. If in this reduced matrix $a_{ij} = 0$ for all $i \geq 2$, then it is already in reduced row echelon form. If not, let j_2 be the smallest number such that $a_{ij_2} \neq 0$ for some $i \geq 2$. Further, let i_2 be the smallest number greater than 2 such that $a_{i_2 j_2} \neq 0$. Note that $j_2 > j_1$. Interchanging the i_2th row and the second row, we may assume that $a_{2j_2} \neq 0$. Then multiplying the second row by $a_{2j_2}^{-1}$, we may assume that $a_{2j_2} = 1$. In turn, adding $-a_{ij_2}$ times the second row to the ith row for each $i \neq 2$, A may have been reduced to a matrix in reduced row echelon form. If not, proceed as before. This process reduces A in to reduced row echelon form after finitely many steps (if worst comes, at the n_{th} step). ♯

Corollary 2.4.16 *Row rank of a matrix is the same as the column rank of the matrix.*

Proof From the Proposition 2.4.6, and the Proposition 2.4.7, row rank and column rank of a matrix are invariant under elementary row operations. From the Proposition 2.4.10(iv), row rank of a matrix in reduced row echelon form is same as its column rank (equal to the number of pivots). Combining this with the Theorem 2.4.15, the result follows. ♯

Definition 2.4.17 Row rank of a matrix A, or equivalently, the column rank of a matrix is called the **rank** of the matrix. The rank of a matrix A is denoted by $r(A)$.

Corollary 2.4.18 *Let A be a $m \times n$ matrix. Then $r(A) + n(A) = n$.*

Proof Since the rank and the nullity remain invariant under elementary row operations, using Theorem 2.4.15, it is sufficient to prove the result for matrices in reduced row echelon form. For a matrix A in reduced row echelon form, $r(A)$ is the number of pivot columns and $n(A)$ is the number of free columns. Clearly, a column is either a pivot column or a free column. ♯

Example 2.4.19 Consider the system of linear equations

$$2x_3 + 3x_4 + 8x_5 = 1.$$

$$2x_1 + 4x_2 + x_3 + 5x_5 = 0.$$

$$x_1 + 2x_2 + x_3 + x_4 + 5x_5 = 2.$$

$$5x_1 + 10x_2 + 6x_3 + 6x_4 + 28x_5 = a.$$

The corresponding coefficient matrix A is

$$A = \begin{bmatrix} 0 & 0 & 2 & 3 & 8 \\ 2 & 4 & 1 & 0 & 5 \\ 1 & 2 & 1 & 1 & 5 \\ 5 & 10 & 6 & 6 & 28 \end{bmatrix},$$

and the augmented matrix A^+ is

$$A^+ = \begin{bmatrix} 0 & 0 & 2 & 3 & 8 & 1 \\ 2 & 4 & 1 & 0 & 5 & 0 \\ 1 & 2 & 1 & 1 & 5 & 2 \\ 5 & 10 & 6 & 6 & 28 & a \end{bmatrix}.$$

We discuss the consistency of the above system of linear equations, and if consistent, we determine a general solution. For the purpose, we reduce the coefficient matrix A, and also the augmented matrix A^+ to reduced row echelon forms simultaneously by using the algorithm described in the above theorem. The 1st column of A is nonzero, and the smallest number i for which $a_{i1} \neq 0$ is 2. Thus, interchanging the 1st and the 2nd rows of A, and of A^+, A is transformed to

$$\begin{bmatrix} 2 & 4 & 1 & 0 & 5 \\ 0 & 0 & 2 & 3 & 8 \\ 1 & 2 & 1 & 1 & 5 \\ 5 & 10 & 6 & 6 & 28 \end{bmatrix},$$

and A^+ is transformed to

$$\begin{bmatrix} 2 & 4 & 1 & 0 & 5 & 0 \\ 0 & 0 & 2 & 3 & 8 & 1 \\ 1 & 2 & 1 & 1 & 5 & 2 \\ 5 & 10 & 6 & 6 & 28 & a \end{bmatrix}.$$

Now, multiplying the 1st row by $\frac{1}{2}$, the matrices are transformed to

$$\begin{bmatrix} 1 & 2 & \frac{1}{2} & 0 & \frac{5}{2} \\ 0 & 0 & 2 & 3 & 8 \\ 1 & 2 & 1 & 1 & 5 \\ 5 & 10 & 6 & 6 & 28 \end{bmatrix},$$

and to

$$\begin{bmatrix} 1 & 2 & \frac{1}{2} & 0 & \frac{5}{2} & 0 \\ 0 & 0 & 2 & 3 & 8 & 1 \\ 1 & 2 & 1 & 1 & 5 & 2 \\ 5 & 10 & 6 & 6 & 28 & a \end{bmatrix}.$$

Further, adding -1 times the 1st row to the 3rd row, and adding -5 times the 1st row to the 4th row, the matrices are transformed to

$$\begin{bmatrix} 1 & 2 & \frac{1}{2} & 0 & \frac{5}{2} \\ 0 & 0 & 2 & 3 & 8 \\ 0 & 0 & \frac{1}{2} & 1 & \frac{5}{2} \\ 0 & 0 & \frac{7}{2} & 6 & \frac{31}{2} \end{bmatrix},$$

and to

$$\begin{bmatrix} 1 & 2 & \frac{1}{2} & 0 & \frac{5}{2} & 0 \\ 0 & 0 & 2 & 3 & 8 & 1 \\ 0 & 0 & \frac{1}{2} & 1 & \frac{5}{2} & 2 \\ 0 & 0 & \frac{7}{2} & 6 & \frac{31}{2} & a \end{bmatrix}.$$

Here, in this transformed matrix, $a_{i2} = 0 \textit{ for all } i \geq 2$. Thus, the 2nd column is a free column. We look at the 3rd column. The 2nd row 3rd column entry $a_{23} = 2 \neq 0$. We divide the 2nd row by 2 to get the pivot entry 1 in 2nd row 3rd column. The matrices, thus, reduce to

$$\begin{bmatrix} 1 & 2 & \frac{1}{2} & 0 & \frac{5}{2} \\ 0 & 0 & 1 & \frac{3}{2} & 4 \\ 0 & 0 & \frac{1}{2} & 1 & \frac{5}{2} \\ 0 & 0 & \frac{7}{2} & 6 & \frac{31}{2} \end{bmatrix},$$

and to

$$\begin{bmatrix} 1 & 2 & \frac{1}{2} & 0 & \frac{5}{2} & 0 \\ 0 & 0 & 1 & \frac{3}{2} & 4 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 1 & \frac{5}{2} & 2 \\ 0 & 0 & \frac{7}{2} & 6 & \frac{31}{2} & a \end{bmatrix}.$$

In turn, to make all other entries in this pivot column 0, we add $-\frac{1}{2}$ times the 2nd row to the 1st row, $-\frac{1}{2}$ times the 2nd row to the 3rd row, and $-\frac{7}{2}$ times the 2nd row to the 4th row. The matrices reduce to

$$\begin{bmatrix} 1 & 2 & 0 & -\frac{3}{4} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{3}{2} & 4 \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{3}{4} & \frac{3}{2} \end{bmatrix},$$

and to

$$\begin{bmatrix} 1 & 2 & 0 & -\frac{3}{4} & \frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{3}{2} & 4 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{7}{4} \\ 0 & 0 & 0 & \frac{3}{4} & \frac{3}{2} & a-\frac{7}{4} \end{bmatrix}.$$

The 3rd row 4th column entry $a_{34} = \frac{1}{4} \neq 0$. We multiply the 3rd row by 4 to get the pivot entry 1 in 3rd row 4th column. Thus, the matrices further reduce to

$$\begin{bmatrix} 1 & 2 & 0 & -\frac{3}{4} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{3}{2} & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & \frac{3}{4} & \frac{3}{2} \end{bmatrix},$$

and to

$$\begin{bmatrix} 1 & 2 & 0 & -\frac{3}{4} & \frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{3}{2} & 4 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & \frac{3}{4} & \frac{3}{2} & a-\frac{7}{4} \end{bmatrix}.$$

In turn, we add $\frac{3}{4}$ times the 3rd row to the 1st row, $-\frac{3}{2}$ times the 3rd row to the 2nd row, and the $\frac{-3}{4}$ times the 3rd row to the 4th row to make the rest of the entries in this pivot column 0. The coefficient matrix A reduces to the following matrix

$$\begin{bmatrix} 1 & 2 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

which is in reduced row echelon form, and the augmented matrix A^+ gets transformed to

$$\begin{bmatrix} 1 & 2 & 0 & 0 & 2 & 5 \\ 0 & 0 & 1 & 0 & 1 & -10 \\ 0 & 0 & 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 0 & 0 & a-7 \end{bmatrix}.$$

Thus, the given system of linear equations is equivalent to a system of linear equations whose coefficient matrix is

$$\begin{bmatrix} 1 & 2 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and the augmented matrix is

$$\begin{bmatrix} 1 & 2 & 0 & 0 & 2 & 5 \\ 0 & 0 & 1 & 0 & 1 & -10 \\ 0 & 0 & 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 0 & 0 & a-7 \end{bmatrix}.$$

In turn, using the discussions and the results above, we have the following: (i) A basis of the row space of A is $\{(1, 2, 0, 0, 2), (0, 0, 1, 0, 1), (0, 0, 0, 1, 2)\}$. The rank $r(A) = 3$.
(ii) Putting the free variable $x_2 = 1$, and the free variable $x_5 = 0$, we get a solution $(-2, 1, 0, 0, 0)$ of the homogeneous part of the system. Further, putting the free variable $x_2 = 0$, and the free variable $x_5 = 1$, we get another solution $(-2, 0, -1, -2, 1)$ of the homogeneous part of the system. The set $\{(-2, 1, 0, 0, 0), (-2, 0, -1, -2, 1)\}$ is a basis of the solution space $N(A)$ of the homogeneous part. A general solution of the homogeneous part of the system is

$$c_1(-2, 1, 0, 0, 0) + c_2(-2, 0, -1, -2, 1),$$

where c_1, c_2 are arbitrary constants.
(iii) The nonhomogeneous system is consistent if and only if $3 = r(A) = r(A^+)$, or equivalently, $a = 7$. Then, giving the value $x_2 = 1$, and $x_5 = 0$ of the free variables, in the nonhomogeneous system, we get a particular solution $(3, 1, -10, 7, 0)$. Thus, a general solution of the nonhomogeneous part is

$$(3, 1, -10, 7, 0) + c_1(-2, 1, 0, 0, 0) + c_2(-2, 0, -1, -2, 1),$$

where c_1, c_2 are arbitrary constants.

Definition 2.4.20 A square matrix E obtained by applying elementary row operations on identity matrix is called an **elementary matrix**.

Example 2.4.21 The matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is an elementary matrix which is obtained by multiplying the 3rd row of the identity matrix I_4 by 3. The matrix

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is an elementary matrix which is obtained by interchanging the 1st row and the 3rd row of the identity matrix I_4. Again, the matrix

$$\begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is also an elementary matrix which is obtained by adding 3 times the 3rd row of the identity matrix I_4 to its 1st row.

τ_{ij} denotes the elementary matrix which is obtained by interchanging the i_{th} row and the j_{th} row of identity matrix. Thus,

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \tau_{13}.$$

The elementary matrix which is obtained by adding the λ times the j_{th} row of the identity matrix to its i_{th} row is denoted by E_{ij}^{λ}. Indeed, E_{ij}^{λ} is the matrix all of whose diagonal entries are 1, the i_{th} row j_{th} column entry is λ, and the rest of entries are 0. Thus,

$$\begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = E_{13}^{3}$$

The matrices E_{ij}^{λ} are called the **transvections**.

It can be easily observed that the effect of multiplying an elementary matrix E from left (right) to a matrix A is applying the elementary row (column) operation on A which was used to get the matrix E from the identity matrix. Thus, $\tau_{ij}A$ is the matrix obtained by interchanging i_{th} row and j_{th} row of A, and $E_{ij}^{\lambda}A$ is the matrix obtained by adding λ times the j_{th} row of A to its i_{th} row. It is straightforward, in particular, to verify the following relations, called the **Steinberg relations**, among the transvections in $M_n(F)$.

(i) $E_{ij}^{\lambda} \cdot E_{ij}^{\mu} = E_{ij}^{\lambda+\mu}, i \neq j$. In particular, $E_{ij}^{\lambda} \cdot E_{ij}^{-\lambda} = E_{ij}^{0} = I_n$. Thus, E_{ij}^{λ} is invertible, and its inverse is $E_{ij}^{-\lambda}$.
(ii) For $i \neq l, j \neq k,\ E_{ij}^{\lambda}$ and E_{kl}^{μ} commute.
(iii) For $i \neq l$, $(E_{ij}^{\lambda}E_{jl}^{\mu}E_{ij}^{-\lambda}E_{jl}^{-\mu}) = E_{il}^{\lambda\mu}$.
(iv) For $j \neq k$, $(E_{ij}^{\lambda}E_{ki}^{\mu}E_{ij}^{-\lambda}E_{ki}^{-\mu}) = E_{jk}^{-\mu\lambda}$.

Proposition 2.4.22 *Let A be a $m \times n$ matrix. Then, we can find a nonsingular $m \times m$ matrix P such that PA is a matrix in reduced row echelon form. In particular, a square matrix A is nonsingular if and only if its reduced row echelon form PA is the identity matrix.*

Proof Applying an elementary row operation on A is equivalent to multiply A from left by an elementary matrix. Since every matrix can be reduced to a matrix in reduced row echelon form (Theorem 2.4.15), multiplying A successively by elementary matrices from left we arrive at matrix in reduced row echelon form. Since elementary matrices are nonsingular, and product of nonsingular matrices are nonsingular, we get a nonsingular matrix P such that PA is a matrix in reduced row echelon form. Since P is nonsingular, A is nonsingular if and only if PA is nonsingular. From the Proposition 2.4.14, A is nonsingular if and only if PA is the identity matrix. ♯

The above discussion and the results give an algorithm to determine a nonsingular matrix P such that PA is a reduced row echelon matrix. In particular, it gives an algorithm to check if a square matrix A is invertible, and if so, to find the inverse of A. We further illustrate the algorithm by means of examples.

Example 2.4.23 Consider the matrix

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

Using the elementary row operations, we transform the matrix A in to a matrix in reduced row echelon form, and simultaneously find a nonsingular matrix P such that PA is a matrix in reduced row echelon form. We start with the pair

$$\left[I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, A = \begin{bmatrix} 0 & 0 & 3 & 1 & 2 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 2 & 1 & -1 & 1 \\ 0 & 1 & 1 & -\frac{1}{3} & 0 \end{bmatrix} \right].$$

There is no nonzero entry in the first column of A, and so no pivot will appear in the first column. We leave and move to the second column. The first nonzero entry in the second column of A is 1, and it is in the second row. We interchange the first row R_1 and the second row R_2 in the pair of matrices. The pair, thus, gets transformed to the pair (E_1, A_1) given by

$$\left[E_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, A_1 = \begin{bmatrix} 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 2 & 1 & -1 & 1 \\ 0 & 1 & 1 & -\frac{1}{3} & 0 \end{bmatrix} \right]$$

(note that $E_1 A = A_1$). The entry 1 in the first row and second column of A_1 is the pivot entry. To make the rest of the entries in this pivot column 0, we replace R_3 by $R_3 - 2R_1$, and then R_4 by $R_4 - R_1$. In turn, the pair (E_1, A_1) gets transformed to the pair (E_2, A_2) given by

$$\left[E_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & -3 & -1 & 1 \\ 0 & 0 & -1 & -\frac{1}{3} & 0 \end{bmatrix} \right]$$

(Again note that $E_2A_1 = E_2E_1A = A_2$). The second row third column entry is 3 which is nonzero. We replace R_2 by $\frac{R_2}{3}$ to make it a pivot entry 1, and in turn, we replace R_1 by $R_1 - 2R_2$, R_3 by $R_3 + 3R_2$, and R_4 by $R_4 + R_2$ to make all the rest of the entries in this pivot column 0. Thus, the pair (E_2, A_2) is transformed to the pair (E_3, A_3) given by

$$\left[E_3 = \begin{bmatrix} -\frac{2}{3} & 1 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ \frac{1}{3} & -1 & 0 & 1 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 1 & 0 & -\frac{2}{3} & -\frac{4}{3} \\ 0 & 0 & 1 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & \frac{2}{3} \end{bmatrix} \right]$$

(Again, note that $E_3A_2 = A_3$). Since the 3rd row 4th column, and 4th row 4th column entries are 0, there is no pivot in the 4th column, it is a free column. We go to the 5th column. The 3rd row 5th column entry is 3 which is nonzero. We replace R_3 by $\frac{1}{3}R_3$ to make the 3rd row 5th column entry a pivot entry 1, and then replace R_1 by $R_1 + \frac{4}{3}R_3$, R_2 by $R_2 - \frac{2}{3}R_3$, and R_4 by $R_4 - \frac{2}{3}R_3$. Thus, the pair (E_3, A_3) is transformed to the pair (E_4, A_4) given by

$$\left[E_4 = \begin{bmatrix} -\frac{2}{9} & \frac{1}{9} & \frac{4}{9} & 0 \\ \frac{1}{9} & \frac{4}{9} & -\frac{2}{9} & 0 \\ \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{9} & \frac{5}{9} & -\frac{2}{9} & 1 \end{bmatrix}, A_4 = \begin{bmatrix} 0 & 1 & 0 & -\frac{2}{3} & 0 \\ 0 & 0 & 1 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right],$$

where A_4 is in reduced row echelon form, and $P = E_4$ is an invertible matrix such that $PA = A_4$ is in reduced row echelon form.

Example 2.4.24 Consider the matrix A given by

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

We apply the following elementary row operations in succession.

(i) Interchange R_1 and R_2,

(ii) replace R_4 by $R_4 - R_1$, R_3 by $R_3 - 2R_2$ and R_4 by $R_4 - R_2$,

(iii) replace R_3 by $-\frac{1}{3}R_3$, R_1 by $R_1 - 2R_3$, R_2 by $R_2 - 2R_3$ and R_4 by $R_4 + 3R_3$.

on A, and also on I_4. Then, A reduces to the reduced row echelon form

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

and I_4 reduces to

$$P = \begin{bmatrix} -\frac{4}{3} & 1 & \frac{2}{3} & 0 \\ -\frac{1}{3} & 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & -\frac{1}{3} & 0 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

Thus, PA is in the row echelon form given above.

Example 2.4.25 Consider the 3×3 matrix A given by

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$$

If we use the method of the above example, then A reduces to the identity matrix I_3, and I_3 reduces to

$$P = \begin{bmatrix} 3 & -\frac{5}{2} & \frac{1}{2} \\ -3 & 4 & -1 \\ 1 & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}$$

Thus, A is invertible, and $PA = I_3$. Hence P is the inverse of A.

2.5 LU Factorization

If the coefficient matrix of a system of linear equations is upper triangular square matrix U with nonzero diagonal entries, then the solution is easily obtained by inspection. For example, if a system of linear equations is given by the matrix equation $U\overline{x}^t = \overline{b}^t$, where

$$U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 9 \end{bmatrix},$$

and $\overline{b} = (b_1, b_2, b_3)$, then, evidently, the solution is $(18b_1 - \frac{b_2}{2} + \frac{4b_3}{9}, \frac{3b_2 - b_3}{2}, \frac{b_3}{9})$. Similarly, it is also easy to solve a system of linear equations whose coefficient matrix is lower triangular square matrix with nonzero diagonal entries. Further, suppose the coefficient matrix A is invertible, and it is expressed as $A = LU$, where L is a lower triangular matrix, and U is an upper triangular matrix. Then, we first find the solution $\overline{v}$ of $U\overline{y}^t = \overline{b}^t$, and then the solution $\overline{u}$ of $U\overline{x}^t = \overline{v}^t$. Clearly, $\overline{u}$ is the solution of $A\overline{x}^t = \overline{b}^t$.

The above discussion prompts us to look at the problem of factorizing an invertible matrix A as a product LU of a lower triangular matrix L and an upper triangular matrix U. This, in general, is not possible.

Example 2.5.1 Suppose that

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \cdot \begin{bmatrix} u & v \\ 0 & w \end{bmatrix}.$$

Then $au = 0$, $av = 1$, $bu = 1$. This, however, is impossible. This shows that the invertible matrix

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

cannot be expressed as product of a lower triangular and an upper triangular matrix. Observe that the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 3 \\ 0 & 2 & 9 \end{bmatrix}$$

is also not expressible as product of a lower triangular and an upper triangular matrix.

The reason behind the impossibility of expressing the above matrices as product of lower and upper triangular matrices is while reducing these matrices in to reduced row echelon forms, we are forced either to interchange rows, or to add a nonzero multiple of a k_{th} row to l_{th} row for some $k > l$. Equivalently, we need to multiply from left by a corresponding elementary matrix τ_{ij}, or by a corresponding matrix E_{kl}^{λ}. Obviously, these matrices are not lower triangular matrices. Indeed, if, while reducing A in to reduced row echelon form, elementary row operations of the above type are not needed, then we can find a lower triangular matrix P with diagonal entries 1 so that PA is upper triangular. In turn, $A = LU$, where $L = P^{-1}$. We illustrate it by means of examples.

Example 2.5.2 Consider the matrix A given by

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix},$$

and the system of linear equations given by the matrix equation

$$A\overline{x}^t = [1, 2, 3]^t.$$

Adding -1 times the 1st row of A to the 2nd row, and then adding -1 times the 1st row to the 3rd row, or equivalently, multiplying the matrix $E_{13}^{-1}E_{12}^{-1}$ to A from left, we obtain that

$$E_{13}^{-1}E_{12}^{-1}A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 3 & 8 \end{bmatrix}.$$

Again, adding -3 times the 2_{nd} row of the above matrix to its 3_{rd} row, or equivalently, multiplying E_{23}^{-3} to $E_{13}^{-1}E_{12}^{-1}A$ from left, we obtain that $E_{23}^{-3}E_{13}^{-1}E_{12}^{-1}A$ is the upper triangular matrix U given by

$$U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}.$$

Thus, $A = LU$, where $L = E_{12}^{1}E_{13}^{1}E_{23}^{3}$ is the lower triangular matrix given by

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 3 & 1 \end{bmatrix}$$

Now, to find solution of $A\overline{x}^t = [1, 2, 3]^t$, we first find the solution of $L\overline{y}^t = [1, 2, 3]^t$. Equating the corresponding entries of both sides, $y_1 = 1, y_1 + y_2 = 2,$ *and* $y_1 + 3y_2 + y_3 = 3$. This gives the solution $[1, 1, -1]^t$ of $L\overline{y}^t = [1, 2, 3]^t$. Finally, we find the solution of $U\overline{x}^t = [1, 1, -1]^t$ to get the solution of the original equation $A\overline{x}^t = [1, 2, 3]^t$. Equating the entries of both sides in the equation $U\overline{x}^t = [1, 1, -1]^t$, we get that $2x_3 = -1, x_2 + 2x_3 = 1,$ *and* $x_1 + x_2 + x_3 = 1$. Evidently, $x_3 = \frac{-1}{2}, x_2 = 2,$ *and* $x_1 = \frac{-1}{2}$.

2.6 Equivalence of Matrices, Normal Form

Definition 2.6.1 Two $m \times n$ matrices A and B with entries in a field F are said to be equivalent if there exists a nonsingular $m \times m$ matrix P, and a nonsingular $n \times n$ matrix Q such that $A = PBQ$.

Clearly, the relation of being equivalent to is an equivalence relation on $M_{mn}(F)$. We determine a unique representative of each equivalence class of equivalent matrices.

Definition 2.6.2 A $m \times n$ matrix A is said to be in **normal form** if there is $r \leq min(m, n)$ such that

$$A = \begin{bmatrix} I_r & O_{r\ n-r} \\ O_{m-r\ r} & O_{m-r\ n-r} \end{bmatrix},$$

where $O_{m\ n}$ denote the zero $m \times n$ matrix.

Theorem 2.6.3 *Every $m \times n$ matrix is equivalent to a unique matrix in normal form.*

Proof Applying an elementary row operation on a matrix A is equivalent to multiply A from left by an elementary matrix, and applying an elementary column operation is equivalent to multiply matrix A from right by an elementary matrix. Since all elementary matrices are nonsingular, and product of nonsingular matrices are nonsingular, it is sufficient to show that every matrix can be reduced to a matrix in normal form with the help of elementary row, and elementary column operations. The proof of this fact is by the induction on $max(m, n)$, where m is the number of rows and n the number of columns. If $max(m, n) = 1$, then $m = 1 = n$, and $A = [a_{11}]$ is 1×1 matrix. If $A = [0]$, then it is already in normal form. If $a_{11} \neq 0$, then multiplying the row by a_{11}^{-1}, we reduce it to the normal form [1]. Assume that the result is true for all $r \times s$ matrices with $max(r, s) < max(m, n)$. Let $A = [a_{ij}]$ be a $m \times n$ times matrix. If $A = O_{m\,n}$, then it is already in normal form, and there is nothing to do. Suppose that $A \neq O_{m\,n}$. Suppose that $a_{kl} \neq 0$. Interchanging 1_{st} row and kth row, and then interchanging 1_{st} column and the lth column, we may suppose that $a_{11} \neq 0$, and then multiplying the 1_{st} row by a_{11}^{-1}, we may further suppose that $a_{11} = 1$. After this we add $-a_{1j}$ times the first column to the jth column, and then $-a_{i1}$ times the first row to the ith row for all $i \neq 1 \neq j$. This reduces the matrix A into the form

$$\begin{bmatrix} I_1 & O_{1\,n-1} \\ O_{m-1\,1} & B \end{bmatrix},$$

where B is $m - 1 \times n - 1$ matrix. This also gives us a nonsingular $m \times m$ matrix C, and a $n \times n$ nonsingular matrix D such that

$$CAD = \begin{bmatrix} I_1 & O_{1\,n-1} \\ O_{m-1\,1} & B \end{bmatrix}.$$

By the induction hypothesis there is a $m - 1 \times m - 1$ nonsingular matrix C', and there is a nonsingular $n - 1 \times n - 1$ matrix D' such that

$$C'BD' = \begin{bmatrix} I_{r-1} & O_{r-1\,n-r} \\ O_{m-r\,r-1} & O_{m-r\,n-r} \end{bmatrix}$$

Take

$$C'' = \begin{bmatrix} I_1 & O_{1\,n-1} \\ O_{m-1\,1} & C' \end{bmatrix},$$

and

$$D'' = \begin{bmatrix} I_1 & O_{1\,n-1} \\ O_{m-1\,1} & D' \end{bmatrix}.$$

Then C'' and D'' are nonsingular. In fact,

$$(C'')^{-1} = \begin{bmatrix} I_1 & O_{1\,n-1} \\ O_{m-1\,1} & (C')^{-1} \end{bmatrix}$$

(Use block multiplication to show this). Again, using block multiplication, we find that

$$C'' \cdot \begin{bmatrix} I_1 & O_{1\ n-1} \\ O_{m-1\ 1} & B \end{bmatrix} \cdot D'' = \begin{bmatrix} I_1 & O_{1\ n-1} \\ O_{m-1\ 1} & C'BD' \end{bmatrix} = \begin{bmatrix} I_r & O_{r\ n-r} \\ O_{m-r\ r} & O_{m-r\ n-r} \end{bmatrix}$$

Take $P = C \cdot C''$, and $Q = D \cdot D''$. Then P is nonsingular $m \times m$ matrix, and Q a nonsingular $n \times n$ matrix such that

$$PAQ = \begin{bmatrix} I_r & O_{r\ n-r} \\ O_{m-r\ r} & O_{m-r\ n-r} \end{bmatrix}$$

is in normal form. Finally,

$$\begin{bmatrix} I_r & O_{r\ n-r} \\ O_{m-r\ r} & O_{m-r\ n-r} \end{bmatrix}$$

is equivalent to

$$\begin{bmatrix} I_s & O_{s\ n-s} \\ O_{m-s\ s} & O_{m-s\ n-s} \end{bmatrix}$$

if and only if $r = s$, for one can be obtained from the other using elementary operations if and only if $r = s$. ♯

Corollary 2.6.4 *There are* $min(m, n) + 1$ *equivalence classes of equivalent matrices in* $M_{mn}(F)$.

Proof There are $min(m, n) + 1$ matrices in $M_{mn}(F)$ which are in normal form. ♯

Corollary 2.6.5 *Two matrices A and B are equivalent if and only if they have same rank.*

Proof Since under elementary operations rank of the matrices do not change and rank of the matrix

$$\begin{bmatrix} I_r & O_{r\ n-r} \\ O_{m-r\ r} & O_{m-r\ n-r} \end{bmatrix}$$

is r, the result follows. ♯

Corollary 2.6.6 *All nonsingular matrices in* $M_n(F)$ *are equivalent to* I_n. *The group* $GL(n, F)$ *is a single complete equivalence class of equivalent matrices.* ♯

Proof Let A be a $n \times n$ matrix which is nonsingular. Then there are nonsingular matrices P and Q such that PAQ is in normal form. Clearly, then PAQ is also nonsingular. The result follows if we observe that a matrix in normal form is nonsingular if and only if it is the identity matrix. ♯

Corollary 2.6.7 *The group* $GL(n, F)$ *is generated by elementary matrices. Indeed, every element of* $GL(n, F)$ *is product of elementary matrices.*

Proof All elementary matrices are nonsingular, and so they belong to $GL(n, F)$. Further, given any matrix $A \in GL(n, F)$, there are nonsingular matrices P and Q which are product of elementary matrices such that $PAQ = I_n$. But, then $A = P^{-1}Q^{-1}$. Since inverse of an elementary matrix is an elementary matrix, P^{-1} and Q^{-1} are product of elementary matrices. This shows that A is product of elementary matrices. ♯

Remark 2.6.8 The matrices $\{E_{ij}^{\lambda} \mid i \neq j, \lambda \in F^{\star}\}$ do not generate $GL(n, F)$ (verify).

Remark 2.6.9 The proof of the Theorem 2.6.3 gives us a method by which

(i) we can reduce a matrix A into normal form,
(ii) we can find nonsingular matrices P and Q such that PAQ is in normal form, and
(iii) we can determine whether A is nonsingular, and then we can find its inverse also.

Following two examples illustrates the algorithm.

Example 2.6.10 Let A be a $m \times n$ matrix. To find nonsingular matrices P and Q such that PAQ is in normal form, we proceed as follows: We start with a row with three columns. The first column I_m, the second A, and the third column I_n. Then we try to reduce the matrix A in to normal form by successive elementary row and elementary column operations. Whenever we perform a row operation on A, apply the same operation to the matrix in the first column, and keep the matrix in the third column as it is, and if we perform a column operation on A, then we perform the same operation on the matrix in the third column, and keep the matrix in the first column as it is. Then as the matrix A reduces to a matrix in normal form, the matrix in the first column reduces to the required matrix P, and the matrix in the third column reduces to the required matrix Q. Consider, for example, the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}.$$

Let R_i denote the ith row, and C_j denote the jth column. We start with a row

$$\left[\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right].$$

Replacing R_2 by $R_2 - 2R_1$, and R_3 by $R_3 - R_1$, we transform the above row to the row

$$\left[\begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right].$$

Next, replacing C_2 by $C_2 - C_1$, and C_3 by $C_3 - C_1$, we get the transformed row as

$$\left[\begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right].$$

Interchanging R_2 and R_4, and then replacing R_4 by $R_4 + 2R_2$, it reduces to

$$\left[\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ -2 & 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right].$$

Replacing C_3 by $C_3 - 2C_2$, we transform it to

$$\left[\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ -2 & 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}\right].$$

Finally, replacing R_3 by $-R_3$, and then R_4 by $R_4 - 3R_3$, we transform it to

$$\left[\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ -5 & 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}\right].$$

Thus, A reduces to the normal form

$$\begin{bmatrix} I_3 \\ O_{1\ 3} \end{bmatrix}.$$

Further, the required nonsingular matrices P and Q are given by

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ -5 & 1 & 3 & 2 \end{bmatrix},$$

and

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

2.7 Congruent Reduction of Symmetric Matrices

Definition 2.7.1 A square matrix A is said to be **congruent** to a matrix B if there is an invertible matrix P such that $PAP^t = B$.

Observe that if A is symmetric, then PAP^t is also symmetric.

Theorem 2.7.2 *Every symmetric matrix A with entries in a field F of characteristic different from 2 is congruent to a diagonal matrix.*

Proof The proof is algorithmic. Let us recall that applying an elementary row operation on A is equivalent to multiply from left the corresponding elementary matrix E, and applying the same type of elementary column operation on A is equivalent to multiply the matrix A from right by the elementary matrix E^t (note that if we apply an elementary row operation on the identity matrix and take its transpose, then it is the same as apply the same elementary column operation on the identity matrix). Thus, it is sufficient to show that a symmetric matrix with entries in a field F of characteristic different from 2 can be reduced to a diagonal matrix by applying successively elementary row followed by the same type of elementary column operations. Let A be a symmetric matrix with entries in F, where characteristic of F if different from 2. If $A = 0$, then there is nothing to do. Suppose that $A \neq 0$. We may suppose that $a_{11} \neq 0$, for if not, suppose that $a_{ij} = a_{ji} \neq 0$, then adding the ith row to the first row, and then adding the ith column to the first column the first row first column entry becomes $2a_{ij} \neq 0$ (note that the characteristic $F \neq 2$). Then, for each $i \neq 1$, adding $-a_{i1}a_{11}^{-1}$ times the first row to the ith row, and $-a_{i1}a_{11}^{-1}$ times the first column to the ith column, we reduce the matrix to a symmetric matrix matrix in which all entries in the first row (and so also in the first column) except a_{11} is 0. Now, if $a_{ij} = 0$ for all $i, j \geq 2$, we have reduced it to a diagonal matrix. If not, using the previous argument, we may take $a_{22} \neq 0$, and then for $i \neq 2$ reduce all the entries $a_{i2} = a_{2i} = 0$. Proceeding inductively we reduce the matrix A to a diagonal matrix. ♯

Taking $Q = P^{-1}$, we get the following corollary.

Corollary 2.7.3 *Every symmetric matrix A with entries in a field of characteristic different from 2 can be decomposed as $A = QDQ^t$, where Q is an invertible matrix, and D is a diagonal matrix.* ♯

Remark 2.7.4 The theorem does not hold over a field of characteristic 2. Consider the matrix

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Suppose that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix}.$$

Equating the corresponding entries $p = ba + ab,\ q = dc + cd,\ da + cb = 0 = bc + ad$. Since the field is of the characteristic 2, $p = 0 = q$. In turn,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

But, then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is singular.

We illustrate the algorithm of congruent reduction by means of an example.

Example 2.7.5 Let A be a symmetric $n \times n$ matrix. To find a nonsingular matrix P such that P^tAP is a diagonal matrix, we proceed as follows: We start with a row with 3 columns, the first column I_n, the second column A, and the third column I_n. We reduce the matrix A in to a diagonal form by successive elementary row and corresponding elementary column operations as described in the above theorem. Whenever we apply an elementary row operation on A, we apply the same operation on the matrix in the first column, and keep the matrix in third column as it is, and whenever we apply elementary column operation we apply the same operation on the matrix in the third column, and keep the first column as it is. In this process as soon as A reduces to a diagonal matrix, the first column reduces to P, and the third column, then will be P^t. Further, PAP^t is a diagonal matrix. Consider, for example, the matrix

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$$

and the triple

$$\begin{bmatrix} I_3 & A & I_3 \end{bmatrix}$$

If we apply the following elementary operations

1. $R_1 \longrightarrow R_1 + R_2$,
2. $C_1 \longrightarrow C_1 + C_2$,
3. $R_2 \longrightarrow R_2 - \frac{1}{2}R_1$,
4. $C_2 \longrightarrow C_2 - \frac{1}{2}C_1$,
5. $R_3 \longrightarrow R_3 - \frac{3}{2}R_1$,
6. $C_3 \longrightarrow C_3 - \frac{3}{2}C_1$,

7. $R_3 \longrightarrow R_3 - R_2$,
8. $C_3 \longrightarrow C_3 - C_2$,

successively, on the triple

$$\left(I_3\ A\ I_3\right),$$

then the triple of matrices reduce to the triple

$$\left[\begin{bmatrix} 1 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -1 & -2 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & -1 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & 1 \end{bmatrix}\right].$$

Thus, A is congruent to $diag(2, -\frac{1}{2}, -4)$, and P is the matrix

$$\begin{bmatrix} 1 & 1 & -0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -1 & -2 & 1 \end{bmatrix}.$$

Further, take $L = P^{-1}$ and $D = diag(2, -\frac{1}{2}, -4)$, then $A = LDL^t$. Note that L is not a lower triangular matrix. However, if we consider the matrix

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$$

with the triple

$$\left(I_3\ A\ I_3\right)$$

of matrices and apply the following elementary operations on each member of the triple to reduce A to a diagonal matrix.

1. $R_2 \longrightarrow R_2 - R_1$ and $R_3 \longrightarrow R_3 - 2R_1$,
2. $C_2 \longrightarrow C_2 - C_1$ and $C_3 \longrightarrow C_3 - 2C_1$,
3. $R_3 \longrightarrow R_3 - R_2$,
4. $C_3 \longrightarrow C_3 - C_2$.

Then the triple of matrices reduce to the triple

$$\left(\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}\right)$$

Thus, A is congruent to $diag(1, -1, -3)$ and P is the matrix

$$\begin{bmatrix} 1 & 0 & -0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$$

Further, take $L = P^{-1}$ and $D = diag(2, -\frac{1}{2}, -4)$, then $A = LDL^t$. Note that in this case P and L are lower triangular matrices.

Example 2.7.6 Consider the symmetric matrix

$$A = \begin{bmatrix} 3 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 3 \end{bmatrix}$$

with the triple

$$\begin{pmatrix} I_3 & A & I_3 \end{pmatrix}$$

of matrices and apply the following elementary operations on each member of the triple to reduce A to a diagonal matrix.

1. $R_3 \longrightarrow R_3 + \frac{1}{3}R_1$ and
2. $C_3 \longrightarrow C_3 + \frac{1}{3}C_1$.

Then the triple of matrices reduce to the triple

$$\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{3} & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{8}{3} \end{bmatrix} \begin{bmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)$$

Here again, P is a lower triangular matrix and the diagonal matrix D has all diagonal entries positive. As such, if we take $L = P^{-1}\sqrt{D}$, where $\sqrt{D} = Diag(\sqrt{3}, 1, \sqrt{\frac{8}{3}}$, then $A = LL^t$. Later we shall describe those symmetric matrices which can be expressed as LL^t, where L is a lower triangular matrix.

Exercises

2.7.1 Give two bases of the vector space $Mnm(F)$ of $n \times m$ matrices with entries in a field F over the field F.

2.7.2 Find a basis, and so also the dimension of the vector space $S_n(F)$ of $n \times n$ symmetric matrices with entries in a field F.

2.7.3 Let F be a field of characteristic different from 2. Find a basis, and so also the dimension of the vector space $SS_n(F)$ of $n \times n$ skew symmetric matrices with entries in a field F. Do the same for fields of characteristic 2. Are they same?

2.7.4 Let A be a $n \times m$ matrix. Consider the subset $W = \{B \in M_{mp} \mid AB = 0_{np}\}$ of M_{mp}. Show that W is a subspace of M_{mp}. Further, show that the dimension of W is $pn(A)$, where $n(A)$ denotes the nullity of A.

2.7.5 Show that every square matrix A with entries in a field F of characteristic different from 2 is uniquely expressible as sum of a symmetric matrix, and a skew symmetric matrix. Deduce that vector space $M_n(F)$ is direct $S_n(F) \oplus SS_n(F)$.
Hint. $A = \frac{A+A^t}{2} + \frac{A-A^t}{2}$.

2.7.6 Find a basis, and so also the dimension of the vector space $UT_n(F)$ of upper triangular matrices over F.

2.7.7 The sum of the diagonal entries of a square matrix A is called the **Trace** of A, and it is denoted by $Tr(A)$. Let $sl(n, F)$ denote the set of $n \times n$ matrices with trace 0. Show that $sl(n, F)$ is a vector space with respect the addition of matrices and multiplication by scalars. Find a basis of $sl(n, F)$, and so also its dimension.

2.7.8 Let A and B be square $n \times n$ matrices. Show that $Tr(AB - BA) = 0$. Deduce that $AB - BA$ is never identity matrix. Show by means of an example that it may be a nonsingular diagonal matrix.

2.7.9 Show by means of an example that AA^t need not be same as A^tA.

2.7.10 Consider the co-diagonal $n \times n$ matrix $\Gamma_n = [a_{ij}]$, where $a_{ij} = 1$ if $i + j = n + 1$, and $a_{ij} = 0$, otherwise. Show that Γ_n is symmetric and $\Gamma_n^2 = I_n$. What is the matrix $\Gamma_n A \Gamma_n$.

2.7.11 Describe all 2×2 matrices A such that $A^2 = 0_2$.

2.7.12 Let A be a strictly upper (lower) triangular $n \times n$ matrix. Show that $A^n = 0_n$.

2.7.13 Let A be a square $n \times n$ matrix which is nilpotent in the sense that $A^m = 0_n$ for some m. Show that $I_n + A$ is invertible. Show that

$$I_n + A + A^2 + \cdots + A^{m-1}$$

is the inverse of A. Is the converse of this statement true? Support.

2.7.14 Let $A = [a_{ij}]$ be a square $n \times n$ matrix which commutes with e_{12}. Show that $a_{12} = 0 = a_{21}$, and $a_{11} = a_{22}$. Show that a matrix commutes with all e_{ij} if and only if it is a scalar matrix. Show also that the matrices which commute with all transvections are precisely scalar matrices. Deduce that the center $Z(GL(n, F))$ is precisely $\{aI_n \mid a \in F^\star\}$.

2.7.15 Find a basis, and so also the dimension of the subspaces of $\mathbb{R}^4$ generated by the following subsets:

(i) $\{(1, 0, 2, 1), (2, 1, 3, 2), (7, 4, 9, 5), (1, 5, 6, 1)\}$,

(ii) $\{(1, 1, 1, 1), (1, 0, 2, 3), (1, 0, 4, 9), (1, 0, 8, 27)\}$.

2.7.16 Reduce the following matrices in to reduced row echelon form. Find the bases of their row spaces, column spaces, and Null spaces. Find their rank, and the nullities. Further, for each of the matrices A, find an invertible matrix P such that PA is a reduced row echelon form of A.

$$\begin{bmatrix} 0 & 0 & 3 & -3 & -3 \\ 2 & 4 & 3 & 3 & 1 \\ 2 & 4 & 3 & 3 & 3 \\ 1 & 2 & 2 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 1 & 0 & -1 & 0 \\ -5 & 1 & 3 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 7 & 11 \\ 3 & 7 & 14 & 25 \\ 4 & 11 & 25 & 50 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$$

2.7.17 Check if the following systems of linear equations are consistent, and if so find their general solutions.

1.

$$\begin{aligned} x_1 + 3x_2 + 4x_3 &= 1. \\ 2x_1 - x_2 + x_3 &= 2. \\ 4x_1 + x_2 - x_3 &= 0. \\ 8x_1 - 3x_2 + x_3 &= 3. \end{aligned}$$

2.

$$\begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 + x_5 &= 2. \\ 2x_1 + 4x_2 + 3x_3 + 3x_4 + x_5 &= 8. \\ 2x_1 + 4x_2 + 4x_3 + 2x_4 + 2x_5 &= 8. \\ x_1 + 2x_2 + 2x_3 + x_4 + 2x_5 &= 2. \end{aligned}$$

3.

$$\begin{aligned} 4x_1 - 15x_2 - 2x_3 - 32x_4 &= -40. \\ x_1 - 2x_2 - 3x_4 &= -4. \\ -3x_1 + 16x_2 + 3x_3 + 38x_4 &= 46. \\ x_1 - 6x_2 - x_3 - 14x_4 &= -17. \end{aligned}$$

2.7.18 Find the value of a, if possible, for which the following system of linear equations is consistent.

$$\begin{aligned} 4x_1 - 15x_2 - 2x_3 - 32x_4 &= -40. \\ x_1 - 2x_2 - 3x_4 &= -4. \\ -3x_1 + 16x_2 + 3x_3 + 38x_4 &= 46. \\ x_1 - 6x_2 - x_3 - 14x_4 &= a. \end{aligned}$$

2.7.19 Check if the matrices in exercise 16 have LU decompositions and if so find their LU decompositions.

2.7.20 Express each of the following symmetric matrices as PDP^t, where P is a nonsingular matrix, and D a diagonal matrix. Which of the matrices are expressible as LDL^t, where L is a lower triangular matrix. Also express them, if possible, as LL^t, where L is a lower triangular matrix.

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 3 \\ 1 & 2 & 1 & 3 \\ 1 & 1 & 3 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 6 & 7 & 8 \\ 3 & 7 & 11 & 12 \\ 4 & 8 & 12 & 16 \end{bmatrix}$$

2.7.21 Find the maximum number of arithmetic operations needed to reduce a 3×3 matrix into reduced row echelon form. Generalize it to $n \times n$ matrices.

2.7.22 Write a program in C-Language to check if a system of linear equations is consistent, and if so to find a general solution.

2.7.23 Write a program in C-Language to check if a matrix A admits LU decomposition, and if so to find it.

2.7.24 Write a program in C-Language to check if a symmetric matrix A admits LL^t decomposition, and if so to find it.

http://www.springer.com/978-981-10-4255-3

Algebra 2
Linear Algebra, Galois Theory, Representation theory, Group extensions and Schur Multiplier
Lal, R.
2017, XVIII, 432 p., Hardcover
ISBN: 978-981-10-4255-3